DOCUMENT RESUME

ED 211 738 CE 031 016

AUTHOR Haney, Walt
TITLE Federal Requirements for the Evaluation of Vocational Education Programs.
INSTITUTION Huron Inst., Cambridge, Mass.
SPONS AGENCY National Inst. of Education (ED), Washington, D.C. Educational Policy and Organization Program.
PUB DATE Aug 81
NOTE 81p.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Educational Assessment; Evaluation; *Evaluation Criteria; Evaluation Methods; *Evaluation Needs; *Federal Legislation; *Federal State Relationship; Outcomes of Education; Postsecondary Education; Program Effectiveness; Secondary Education; State Federal Aid; State Programs; *Vocational Education
IDENTIFIERS Education for All Handicapped Children Act; Title I Evaluation and Reporting System; *Vocational Education Amendments 1976; Vocational Education Study

ABSTRACT

In order to determine how vocational education programs should be evaluated, it is first necessary to define vocational education, to describe how vocational education programs are now being evaluated, to find out how other federal education programs are evaluated, and to suggest what criteria might reasonably be set out in future federal legislation for the evaluation of vocational education. Vocational education can be defined as encompassing some 28,000 institutions providing three types of occupational preparation, in nine general types of program areas, at four educational levels, to meet the needs of at least seven special needs groups. Studies by several researcher and Congressional hearings have demonstrated that vocational education is being evaluated on a variety of criteria, with a confusing mix of data. Federal regulations, however, call for each state vocational education program to be evaluated once every five years in terms of planning and operational processes, student achievement, student employment success, and results of additional services provided to special populations. Title I (Education of Disadvantaged Children) and Public Law 94-142 (Education of the Handicapped) have somewhat different evaluation requirements than the Vocational Education Amendments set forth, raising questions of whether evaluation should be seen as a management and accounting activity for the federal government, or as a diagnostic and planning tool for local schools and districts. It is suggested that the federal role in vocational education evaluation requirements be lessened, inasmuch as the federal government provides less than 10 percent of its support. It is also proposed that evaluation criteria be developed pertaining to program improvement and that evaluation of labor market demands be left to the federal government. (KC)

ED211738

Federal Requirements for the Evaluation of Vocational Education Programs

Walt Haney

The Huron Institute
123 Mt. Auburn St.
Cambridge, Ma. 02138

August, 1981

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Paper prepared for the Vocational Education Study of the National Institute of Education. Views expressed do not, of course, necessarily reflect those of any agency or person other than the author.

CE031016

TABLE OF CONTENTS

page

I. Introduction 1

II. What is Vocational Education 3

III. How Are Vocational Education Programs Now Being Evaluated? . 8

The CRC Study, January 1980 13
Congressional Hearings, September 1980 18
Abt Study, December 1980 23
Wentling's Survey, June 1981 25

IV. How Are Other Federal Education Programs Evaluated? 28

Title I 29
Education for All Handicapped Children 33

V. Federal Legislation: What Criteria Should Be Set Out for the Evaluation of Vocational Education? 37

Purposes of Evaluation 38
Employment Success 42
Student Achievement 48
Services to Special Needs Populations 52
Planning and Operational Processes 55
What Else Can Be Done to Make Evaluation Useful for Program Improvement 60

Evaluation procedures 60
Exemplary program identification 62
Technical assistance 63
Evaluation research into special issues 64
Meeting labor market demands 66

References 70

Appendices: Descriptions of Federal Evaluation Requirements for Education Programs, from Boruch & Cordray (1980)

1. Vocational Education
2. Title I
3. PL-94-142 Education for Handicapped Children

I. Introduction

How should vocational education programs be evaluated? This is the question addressed in this paper. Specifically, we address the question from the perspective of federal vocational education legislation. In order to suggest answers to the question, we first address the following questions, which will provide a basis for our answers:

- What is vocational education?
- How are vocational education programs now being evaluated?
- How are other federal education programs evaluated?
- What criteria might reasonably be set out in future federal legislation for the evaluation of vocational education?

Before embarking on at least brief answers to these questions, this introduction describes the background to this paper.

In the hearings prior to the 1976 Education Amendments by the U.S. Congress, there was much criticism of past evaluation and planning efforts connected with vocational education. As the Interim Report of the NIE Vocational Education Study described it:

> . . . in 1976, Congress heard repeated criticisms of the planning process, and in particular of coordination among parties involved and of use of current data. The 1976 Amendments consequently incorporated a number of provisions intended to improve planning and evaluation procedures.
> (National Institute of Education, 1980, p. 916)

The Interim Report goes on to point out that one of the mechanisms by which Congress sought to improve the degree to which planning was informed by relevant data was by establishing new evaluation requirements. Specifically Section 112(b)(1) of the 1976 amendments stipulated that:

> (A) each State shall, during the five-year period of the State plan, evaluate the effectiveness of each program within the State being assisted with funds available under this Act; and the results of these evaluations shall be used to revise the State's programs, . . .

> (B) each State shall evaluate, . . . each such program within the State which purports to impart entry level job skills according to the extent to which program completers and leavers--
>
> (i) find employment in occupations related to their training, and
>
> (ii) are considered by their employers to be well-trained and prepared for employment. . . .

Two themes are apparent in the language of this section. First is that evaluations of effectiveness should be used to revise and improve vocational programs. Second is the notion that programs which seek to prepare students directly for employment should be evaluated in terms of whether they subsequently do find employment related to their training, and are considered well-trained by their employers. The latter idea was a reflection of a broader theme in the 1976 legislation, namely that vocational education programs should be more closely articulated with labor market conditions and demands for skilled labor. After all, it is of little benefit to anyone for a program to produce individuals skilled in, say, repair of television and radio vacuum tubes, when most such instruments now use transistors.

But if the motivations behind the 1976 vocational education evaluation requirements are easy to understand, it is far less clear exactly what the effects of these requirements have been. Available evidence on this point will be reviewed in section III, but for the moment suffice it to explain that pending reauthorization of federal vocational education in 1982, the Vocational Education Study was mandated to analyze "the means of assessing program quality and effectiveness" (National Institute of Education, 1980, Appendix A In connection with this mandate, the study staff asked us to consider

what criteria might reasonably be specified in federal legislation as bases for evaluating vocational education programs. This paper is one product of that consideration.

II. What Is Vocational Education?

This question seems a reasonable place to begin. The first point to note is that the very phrase "vocational education" connotes two different types of goals. The adjective "vocational" indicates clearly that we are not referring simply to general education, but to education for vocations. Vocation refers, of course, not just to work or labor, but to particular occupational callings. Indeed, the oldest meaning of the word in the English language refers to "action on the part of God in calling a person to exercise some function" (Oxford English Dictionary). At the same time, however, the subject of the phrase is education -- "the process of nourishing or rearing a child or young person" (Ibid.).* The point of this distinction is more than merely semantic, for it is easy to forget that vocational education is not synonymous with vocational training, that the former refers to a particular form of education.

But before attempting to draw out any implications from this distinction, let me address the question of this section at a more practical level. What exactly are the sorts of programs funded under federal vocational educational legislation? The vocational education enterprise is "a congery of different systems and not a single system; it is decentralized and diversified. Its programs of instruction are offered in different kinds of institutions." The

*According to the OED, vocation is a slightly older word in the English language (1426), than education (1540).

number of students enrolled in its wide variety of programs and courses was 17.3 million in 1979 (National Institute of Education, 1980). According to Bottoms (1980), the executive director of the American Vocational Association, "Essentially vocational education provides learning experiences that further the occupational awareness, exploration, preparation, and specialized needs of its clients, typically aged 14-60."

According to the latest report of the National Center for Educational Statistics (NCES) on The Condition of Vocational Education (1980), there are nearly 28,000 different institutions nationwide offering vocational education programs. Although the NCES has described these institutions in terms of some 18 different types, it notes that "even the identification and description of the institutions has required in some instances the use of general terms that mask many significant even fundamental differences, in institutional intent or character" (p. 550). The largest single type of institution providing vocational education with federal support is public comprehensive and vocational high schools, accounting for nearly 16,000 schools.

For this level, namely the secondary level, vocational educators typically describe three major purposes for the enterprise:

1. In-depth exploration of occupations to assist in the career planning process.

2. Development of occupational competencies designed to be recognized for advanced placement in post-secondary programs, and

* The NCES report indicates that at least three types of institution offering vocational programs do not receive federal vocational education funds; namely private secondary schools, private noncollegiate postsecondary schools, and correspondence schools.

3. Development of occupational competencies necessary to enter an occupation.
(Minnesota Department of Education, 1978, p. 361)

Vocational education programs focusing on the latter purpose sometimes are described as occupational vocational education programs, whereas programs serving mainly the former two purposes are called non-occupational in the sense that they are not aimed at directly preparing individuals for occupations.

Substantial numbers of secondary vocational students continue beyond high school for some form of postsecondary schooling. Follow-up data on the high school class of 1972 indicate that at some time within four years out of high school 54% of vocational high school graduates enroll in some form of postsecondary education programs (including both academic and vocational postsecondary programs), as compared with 62.5% of general high school graduates (Woods & Haney, 1981, Appendix D). This trend toward pursuit of postsecondary schooling by secondary vocational graduates appears to have been recognized implicitly by Congress in the 1976 Education Amendments in stipulating that "in no case can pursuit of additional education or training by program completers or leavers be considered negatively in these evaluations" of occupational vocational programs (Section 112(b)(1)). Thus, follow-up studies of secondary vocational graduates usually do (and certainly should) distinguish between individuals who are <u>in</u> the labor force and those who are <u>out</u> of the labor force--the latter referring to individuals who are enrolled in school full-time, are homemakers, and several other categories of individuals as well.

This description only begins to convey some of the complexity of the vocational education enterprise. It is also worth noting that vocational

program offerings are traditionally described in terms of nine program areas:

Agriculture
Distributive
Health
Consumer and homemaking
Occupational home economics
Industrial arts
Office occupations
Technical
Trade and industrial

Nevertheless, when these types are disaggregated, only for occupationally specific programs, instructional programs keyed to nearly 100 different occupations can be identified (Condition of Vocational Education, 1980, pp. 583-586), and these 100 occupationally specific instructional programs are reportedly offered at all four levels of the vocational education enterprise (secondary, postsecondary, adult-short-term, and adult-long-term). Such diversity is all the more impressive when it is noted that enrollments in occupationally specific vocational education programs make up only about 7 million of the estimated 17 million enrollments in vocational education programs overall in 1978-79 (including both occupational and nonoccupational programs).

One final aspect of diversity in the vocational enterprise deserves to be mentioned, namely, the population of individuals intended to be served. As far back as the Smith-Hughes Act in 1917, the target population for federally aided vocational education programs has generally been understood to be high-school aged youth preparing for occupations.* However, the

* Woods & Haney, 1979, provides a more detailed account of the evolution of federal legislative goals in this regard.

Vocational Education Act of 1963 greatly elaborated the definition of the population to be served by vocational education. Specifically, the 1963 Act established the intent for vocational education to meet the needs of four groups: (1) youth in high school; (2) youth with special needs; (3) youth and adults in full-time postsecondary programs preparing to enter the labor market; and (4) youth and adults unemployed or at work, and needing training or retraining to achieve employment stability. The special needs group was defined in the 1963 Act to include "persons who have academic, socioeconomic, or other handicaps that prevent them from succeeding in the regular vocational education program." While the "physically handicapped" were explicitly excluded from the special needs group identified in 1963, it was added as a special needs target group for vocational education by the Education Amendments of 1968. Moreover, the Amendments of 1976 further extended the special needs definition to include women, limited English-speaking persons, and American Indians.

In sum, the vocational education enterprise is tremendously complex and variegated. It encompasses, at last count, some 28,000 institutions providing vocational education instruction. It spans three quite different types of occupational preparation (occupational exploration, preparation for advanced placement in subsequent occupational preparation programs, and preparation for direct entry into occupations). Focusing only on occupational vocational education, we see nine different general types of program areas (spanning nearly 100 different occupations for which vocational programs seek to prepare individuals). Moreover, vocational education programs exist at four different levels of the educational system (secondary,

postsecondary, adult short-term and adult long-term). Finally, and on top of all this, we see that vocational education also is intended to meet the needs of at least seven specifically identified special needs groups.

III. How Are Vocational Education Programs Now Being Evaluated?

Given an enterprise of this magnitude and diversity, how are vocational education programs now being evaluated? Before answering this question, let me describe at least briefly what is meant by evaluation. In its root sense, to evaluate means to determine the value or worth of a thing. Program evaluation as a major social enterprise in the United States has developed only since 1965, when the Elementary and Secondary Education Act was passed by the U.S. Congress with explicit provisions for the evaluation of programs funded under the Act. Many different sorts of activities have taken place in the name of evaluation since then, including needs assessments, cost effectiveness studies, planned variation experiments, program documentation, and studies modeled after literary criticism. Nevertheless, for the purpose of this paper we adopt the definition of evaluation used in a recent Congressionally mandated study of evaluation of federally funded elementary and secondary education programs, as follows:

> Evaluation is defined . . . as a study designed to answer questions about what a program does in the interest of making judgments about the program. The questions often addressed include: Who is served? What services are delivered? At what cost? With what effect?
>
> (Boruch & Cordray, 1980, pp. 1-4)*

* See Chapter 2 of the Boruch & Cordray report for a brief review of some of the diverse meanings attached to the word evaluation by different sorts of observers.

Given this definition, how are such questions being answered with respect to vocational education? Perhaps the easiest place to begin answering this question is with a description of intent. A full description of the evaluation requirements for vocational education as described in the Boruch & Cordray (1980) study is provided in Appendix 1 of this paper. Here, let me note simply that while such requirements have been set out at levels of both federal and state government, and contain clear specifications of organizations and officials to be involved in evaluating vocational education, the purpose of this paper is not to consider such procedural matters, but instead to analyze what criteria might reasonably be set out for evaluating vocational education programs at the local level, or to be more precise vocational education programs receiving federal funds.

Federal regulations currently call for each State Department of Education to evaluate the effectiveness of each such vocational education program, once within the five-year period of the state plan in terms of:

(1) planning and operational processes;

(2) student achievement;

(3) student employment success, and

(4) results of additional services provided to special populations.

The evaluation requirements specified in federal regulations in October 1977 were considerably more extensive than the criteria mentioned in the Educational Amendments of 1976. Thus, before answering the question of how vocational education programs are now being evaluated, let us describe at least briefly how the criteria mentioned in the 1976 legislation were transformed into those set out in the 1977 regulations.

As noted in the introduction, in hearings leading to the passage of the 1976 Amendments, there was considerable criticism of past evaluation and planning with respect to vocational education. Congress was particularly concerned that vocational education program offerings were not responsive to labor market demands and to manpower needs. Thus Congress specified that occupational vocational education programs should be evaluated in terms of which program completers and leavers:

i. find employment in occupations related to their training, (and)

ii. are considered by their employers to be well-trained and prepared for employment.

As Hendrickson (1981) recounts, vocational educators "objected strenuously" to these evaluation requirements. First it was argued that such employment criteria do not acknowledge the broader goals of vocational education programs, and the fact that vocational education is more than simply job training. Second, it was maintained that such job placement and performance criteria in effect would hold vocational educators accountable to unfair standards.

> School systems should be held accountable for imparting certain knowledge and skills and for making certain that there is a good fit between what is taught and the jobs available. However, what jobs students actually take is determined by a host of economic and personal factors beyond the school's control. 'Hold us accountable for employability, but not employment,' is the way many vocational educators phrased the concern.
>
> (Hendrickson, 1981, p. 7)

Third, some people voiced the concern that if vocational education programs were held accountable in terms of job placement and employer satisfaction criteria, that there might be a tendency to avoid accepting individuals into vocational programs who might be harder to place, "such as members of minority groups, the disadvantaged, or women in nontraditional

programs--in short, the very groups Congress most wanted (vocational education) to reach " (Hendrickson, 1981, p. 7).

Apparently at least partly as a result of such objections, regulations issued by the U.S. Office of Education in October 1977 specified a broader approach to evaluating vocational education. As mentioned already, the regulations specified four areas in which evaluation should be conducted (namely, planning and operational processes, student achievement, employment success, and results of additional services provided to special legislations). Details on the exact criteria suggested in the regulations under each of these rubrics are presented in Table 1.

The evaluation regulations were further elaborated in a policy memorandum in April 1979. According to Hendrickson (1981), the evaluation policy memo was issued only after "more than a year and one-half of internal struggle and debate within OE over how prescriptive to be" in setting out evaluation requirements. On one hand, there was a desire to allow states and localities flexibility to organize and carry out evaluations suitable to local conditions and hence to provide more leeway for such evaluations to yield locally relevant information suitable for program improvement (and, of course, at the same time to avoid over-regulation). On the other hand, however, there was a desire to be sufficiently prescriptive so that evaluations from different localities would contain comparable data which could be aggregated to provide a national picture of the status of vocational education.

The 1979 regulations apparently were something of a compromise between these two tendencies, though Hendrickson (1981) suggests that "the specifications in the memorandum flow more from a concern . . . to provide a national picture than from a concern for program improvement" (p. 12).

Table 1

Evaluation Specifications in the Regulations

The State board shall, during the five-year period of the State plan, evaluate in quantitative terms the effectiveness of each formally organized program or project supported by Federal, State, and local funds. These evaluations shall be in terms of:

(a) Planning and operational processes, such as:

(1) Quality and availability of instructional offerings;
(2) Guidance, counseling, and placement and follow-up services;
(3) Capacity and condition of facilities and equipment;
(4) Employer participation in cooperative programs of vocational education;
(5) Teacher/pupil ratios; and
(6) Teacher qualifications.

(b) Results of student achievement as measured, for example, by:

(1) Standard occupational proficiency measures;
(2) Criterion-referenced tests; and
(3) Other examinations of students' skills, knowledge, attitudes, and readiness for entering employment successfully.

(c) Results of student employment success as measured, for example, by:

(1) Rates of employment and unemployment;
(2) Wage rates;
(3) Duration of employment; and
(4) Employer satisfaction with performance of vocational education students as compared with performance of persons who have not had vocational education.

(d) The results of additional services, as measured by the suggested criteria under paragraphs (2), (b), and (c) of this section, that the State provides under the Act to these special populations:

(1) Women;
(2) Members of minority groups;
(3) Handicapped persons;
(4) Disadvantaged persons; and
(5) Persons of limited English-speaking ability.

Source: Federal Register, Vol. 42, No. 101, October 3, 1977, Sec. 104.402.

Whatever the balance of motivations behind it, the 1979 memo set out four general guidelines:

(1) It defined the term "program" which was the entity to be evaluated.

(2) It allowed sampling of programs within states and institutions for purposes of evaluation.

(3) It encouraged cyclical patterns of evaluation so that programs of particular types could be evaluated only once within the five-year cycle of each state's five-year plan, and

(4) It set out the policy that OE would not aggregate into national data summaries either program review data or student achievement results, though it would review and summarize such results to show national trends.

Before describing evidence on how these evaluation requirements appear to be working out in practice, we should explain one aspect of the national effort to evaluate vocational education. In addition to setting out evaluation requirements for states and programs, the 1976 Amendments also directed the USOE to develop a national system for reporting vocational education data. This system, subsequently known as VEDS (Vocational Education Data System) was to be based on information derived from mandated evaluations of vocational education, as well as other information. The main idea behind VEDS was that it would provide a system for acquiring standardized and therefore comparable data from different states and localities which could be aggregated so as to produce a national picture of the vocational education enterprise. The reason for mentioning VEDS here is that reaction to the 1976 evaluation requirements and associated regulations appear to have been influenced strongly by their implementation in connection with VEDS, namely the effort to standardize key aspects of evaluation and data reporting across the vocational education enterprise.

So what then has happened as a result? Four sources of evidence are available with which to address this question:

(1) A January 1979 report, The State of the Art of Vocational Education Evaluation: State Evaluation Procedures and Practices by CRC. Education and Human Development Corporation (Smith and Holt, 1979).

(2) Testimony in September 1980 Congressional Hearings on Current Issues in Vocation Education (U.S. Congress, 1980) including the NCES report, The Condition of Vocational Education, which represents the first public report of data gathered through VEDS, and includes a review of the quality of data describing vocational education (in U.S. Congress, 1980, pp. 543-861).

(3) A December 1980 report, Implementation of the Education Amendments of 1976: A Study of State and Local Compliance and Evaluation Practices, by Abt Associates (Berke et al., 1980), and

(4) A June 1981 Survey of State Evaluation Practices in Vocational Education (Wentling, 1981).

In the following paragraphs, we briefly summarize the main findings of these studies concerning criteria for evaluating vocational education programs. The CRC Study, January 1980. This report was based on interviews, both in person and via the telephone with SEA staff members identified by State Directors of Vocational Education as knowledgeable about evaluation procedures in their states. Interviews were conducted with individuals from all fifty states. The report is said to be based on "information that was available in spring 1978" (p. 2). Findings from the interviews were reported in four major categories: (1) purposes of evaluation, and potential uses of findings; (2) evaluation practices; (3) evaluation procedures; and (4) problems in evaluation.

The CRC study reported that "many states could neither describe nor document how information generated by evaluations are actually used" (pp. 4-5), but nevertheless indicated that major purposes identified by individual

interviewed in the fifty states were as follows (with the numbers of states identifying each purpose shown in parentheses):

- Program improvement at the state, local, or institution level (32);
- Fiscal accountability to state and federal authorities (11);
- Program planning at state, local or particularly institutionsl level (10);
- Compliance with federal or state requirements to evaluate or report on program operations (9); and
- Education accountability (6).

It is somewhat difficult to know precisely what to make of these findings because the CRC report does not make clear exactly how some of the five categories of purpose (e.g. fiscal accountability, education accountability, and compliance) differ. Nevertheless, it seems clear from the CRC findings that the major common purpose cited for vocational education evaluations was program improvement.

CRC considered evaluation practice in terms of how states were preparing to meet the 1976 legislative mandate to evaluate (1) program quality (as indicated by planning and operational processes); and (2) program effectiveness (as indicated by student achievement and employment success, specifically whether former vocational students obtain "employment in an area related to their vocational training" and whether "employers are satisfied with the performance of former vocational students"). States were divided into four major categories with regard to their preparedness to conduct evaluations in each of three areas (program review, student achievement; and follow-up surveys) as of spring 1978:

-states planning to use existing systems without revision;

-states planning to revise existing systems in light of the new law;

-states developing or adopting completely new evaluation systems;

-states which had delayed decision on how they will comply with the new law.

Data presented in the CRC report indicate that only two states (North Carolina and Utah) were planning to use existing evaluation systems in each of the three areas of evaluation. Twenty-four states were indicated to be developing a new evaluation with respect to program review, student achievement, and/or follow-up surveys and for fourteen other states it was indicated that no information was available as of spring 1978 regarding what was planned in one or more of the three areas of evaluation.* Of the three areas concerning which states were surveyed, student achievement appeared to be the one for which it was most unclear what states would do with respect to evaluation.

Other salient findings by CRC regarding evaluation practices were as follows:

-"Forty-two states evaluate secondary programs, thirty-seven states evaluate post-secondary programs, while only seventeen states evaluate adult programs" (p. 14);

-"Approximately 28 SEA's have not developed standards of program 'quality' or 'effectiveness'" (p. 16);

-"Among the state standards which have been developed, most focus on program inputs and processes" (p. 14);

-"Only a few SEA's are currently using student achievement as an indicator of program effectiveness, partly because of a lack of available procedures for determining in a consistent fashion what achievement levels are" (p. 23).

*These numbers do not sum to 50, at least partly as a result of some apparently inexplicable gaps in the CRC data.

Under the rubric of evaluation procedures the CRC study inquired into the state mechanisms for evaluation in the areas of program review, student achievement of compentencies at the end of program, and student employment success. "Program review procedures are the most fully developed evaluation procedures in vocational education" (p. 24). Among procedures found to be used for program review were self-study, on-site visit by an outside state-agency selected team, and feedback mechanisms to verify that recommendations for program improvement have been implemented (p. 24). "Procedures for the measurement of student achivement constitute the least developed component of state evaluation systems. Only one-fifth of the states currently have procedures for the assessment of student achievement; those procedures are mostly in developmental or pilot testing stages" (p. 31). "Student employment success serves as a major criteria for assessing program effectiveness. ... Approximately 37 states currently have operational systems which collect student post-program employment data. However, only 25 percent of the states reported that they conduct employer surveys although so required by law" (pp. 31-32).

In the realm of problems in evaluation the CRC report noted that "many administrators identified difficulties experienced either in implementing evaluations or in planning them for the future year," even though "some state administrators appear to be reluctant to discuss their state's outstanding probems in evaluation with individuals outside the SEA" (p. 35). The specific probems identified in the CRC report are the following:

- Insufficient staff and funds to conduct or monitor evaluations prescribed by the legislation.
- Insufficient staff, who are trained in evaluation and data analysis, to design evaluation procedures and make optimal use of evaluation findings, once they are obtained.

-Lack or insufficient access to [data processing] facilities to handle, store, and retrieve information in a timely fashion.

-Lack of clarity about the intent of the law: what information should be included in the evaluation to comply with the new mandate.

-Inaccuracy in reporting, resulting in an incorrect assessment of program operations (p. 35).

Congressional Hearings, September 1980

A second source of information on the current status of vocational education evaluation is testimony offered in hearings held before the House Committee on Education and Labor in September 1980 (U.S. Congress, 1980). While virtually all of the testimony is at least indirectly relevant to the issue of what criteria might be appropriate for evaluating vocational education programs, certain specific observations are directly pertinent to the current status of vocational education evaluation, and hence it is such observations that we will briefly describe.

Rolf Wulfsberg of the NCES offered testimony concerning the current condition of vocational education and the experience of the NCES in developing VEDS. In describing changes in enrollment between 1977-78 and 1978-79, Wulfsberg noted that "There were 14 states and territories that showed . . . significant deviations from the reporting last year, . . . enrollment changes of over 15 percent in either direction." In such cases NCES inquired into possible causes of such changes and it was found that "in about half the cases . . . the state attributes the most of the change to the reporting of accurate data for what they feel is the first time in certain areas. They attribute this to new definitions and the requirement for student unit records contained to the VEDS" (sic, p. 10). Later in response to a question, Wulfsberg offered the following explanation of the problem of duplicated counts of vocational education enrollments:

> When NCES began the development of VEDS in 1977, many factors were preventing the collection of accurate, comparable, unduplicated enrollments across States;
> - o Some States made no attempt to report unduplicated counts;
> - o Some States reported enrollments in courses, while other States reported program enrollments only;
> - o Some States reported enrollments below the 6th grade, while others reported only on students in 12th grade programs which ended in a specific sequence of courses; and

- o Some States only reported enrollments in programs which were directly receiving Federal funds, while others reported all enrollments in sub-baccalaureate vocational programs.

Thus, duplicated counts represented only one of many factors blocking the achievement of our objective.

In order to obtain a rough idea of the extent to which these problems (particularly duplication) still exist, we compared the number of occupationally specific vocational enrollments in grades 11-12 to the count of all enrollments (vocational or nonvocational) in grades 11-12 on a State by State basis. The results indicate that 20 States/Outlying Areas reported occupationally specific vocational enrollments exceeding 50 percent of the total. While in some cases the high percentage was in part due to students below grade 11 being incorrectly reported as in grades 11-12, duplication and the counting of course enrollments (rather than program enrollment) clearly remain a problem. Because the causes described earlier are all present in varying degrees, we cannot specify how much of the overcount is due to duplication in particular. As the States continue to implement student unit record systems, however, duplication will become less and less of a problem in the next two or three years. (p. 238)

Wulfsberg went on to explain the steps that NCES has undertaken in order to obtain accurate unduplicated counts of enrollments in vocational education programs:

The first step was the definition of rules to be used in assigning a vocational student to one and only one program. The only exception was Consumer and Homemaking and Industrial Arts, where a student might be counted twice -- once in Consumer and Homemaking or Industrial Arts and once in another vocational program.

The second step was the separation of occupationally specific enrollments from other vocational enrollments for reporting purposes. Occupationally specific enrollments specifically exclude, among others, enrollments in Consumer and Homemaking and Industrial Arts which may include duplicated counts. Also excluded are enrollments below grade 11. Since the States have varying policies concerning the reporting of students in grades 10 and below, this makes occupationally specific enrollments much more comparable on a State by State basis than any vocational counts previously available.

From the initial development of VEDS, NCES has stressed the need for the establishment of student unit record systems which, by their nature, produce unduplicated, auditable enrollment data. . . .

> In regulations published on May 9, 1980, NCES listed auditable enrollment data as one criterion for full compliance with VEDS reporting requirements (104 CFR 117-119). NCES has stressed in all negotiations with States the need to establish unit record systems, and most States are committed to having such systems operational by the 1980-81 reporting cycle. All States should have auditable data by 1981-82, according to their remediation plans. (pp. 238-239).

Mr. Wulfsberg also commented on the difficulty of obtaining accurate follow-up data on the employment success of former vocational education students:

> I would like to tell you more about what is happening to the placement side of the picture but I must say that we do not feel that the data on followup that have been provided you in the past are valid. They only include a followup program of completers, and have some leavers--those who are known to have found related employment.
>
> This biases the picture and leaves out a very important part of this population, namely the leavers. It implies by default that program leavers are losses to the program, and certainly such is not the case. (p. 15)

Several witnesses in the September 1980 hearings nevertheless testified that VEDS was helping to remedy past problems. Dr. Gene Bottoms, executive director of the American Vocational Association, testified that:

> While a variety of technical and specific, as well as policy level problems remain, it is clear that VEDS is a significant improvement over past systems. Recognition of and action to implement more reliable local data collection practices is still needed, however. (p. 202)

Robert Taylor, executive director of the National Center for Research in Vocational Education, also offered testimony suggesting that evaluation studies of vocational education prior to the establishment of VEDS tended to be less rigorous (p. 259).

Nevertheless, the VEDS system did not escape criticism. Bottoms, for example, criticized the costs of meeting VEDS requirements and the likely utility of VEDS data for state and local purposes, as follows:

> The cost of fully implementing the . . . VEDS may equal the total increase in federal appropriations for vocational education since 1976. This requirement to meet federal data needs has placed massive data collection responsibilities on states and local school districts. It is unlikely, however, that it will provide much timely and useful data for state and local planning and program development.(p. 213)

The record of the September 1980 hearings also contains a letter from a community college chancellor in California, criticizing VEDS as follows:

> The value of such data, particularly when aggregated at the national level, is very questionable. Such aggregation tends to minimize or ignore significant regional, state, and institutional differences. The difficulties inherent in the collection of such data with the detail inherent to the N.C.E.S. design is particulaly evident in respect to the student and employer follow-up section. I am convinced that this strategy will not produce unbiased, valid or reliable data.
>
> Of at least equal concern is the massive potential financial impact of the system. The cost of VEDS implementation has been estimated in some quarters to exceed fifty (50) million dollars annually. This figure may not even adequately measure the local institutional commitment to this data collection task . . . .
>
> It is our hope that [the House Subcommittee on Elementary and Secondary Education] will address these inadequacies in the design of the current VEDS. Any such system should in fact provide valid and useful data and should, in addition, be of value to the institutions and states obligated to its collection. These necessary modifications should not necessarily require the provision of national aggregate data. If such aggregation was not required, the expense and difficulty of developing standardized national data elements would be eliminated. Without such modification it does not appear that VEDS can serve the purposes for which it was established and will continue to impose severe burdens upon the educational institutions. (pp. 540-541)

In addition to touching on the current status of evaluation and data gathering on vocational education, the 1980 Hearings also touched on the question of what criteria might be most appropriate in evaluating the effectiveness of vocational education programs. In his testimony, for example, Taylor

mentioned several different measures of effectiveness--earnings, employment, job satisfaction, and employer satisfaction. Taylor was asked by Carl Perkins, Chairman of the House Committee on Education and Labor, which of these measures he feels is most important in judging the success of vocational programs. Taylor responded as follows:

> It is extremely difficult to assign a higher weight to any of the criteria for measuring the effectiveness of vocational education. Earnings, employment, job satisfaction and employers satisfaction. In an absolute sense, however, employment is probably the most fundamental measure, followed by earnings. It also follows that individuals who have completed vocational education need to be satisfied with the job and employers need to be satisfied with the performance of vocational graduates. I believe that the four measures tend to be a cluster which are extremely difficult to separate and which in fact do interrelate and support each other. (p. 331)

Similarly, Dr. Henry David, Director of the Congressionally mandated Vocational Education Study, was asked by Congressman Arlen Erdahl, "How do you measure the success of vocational education?" He responded as follows:

> I may sound as if I am hedging, Mr. Erdahl, but if you will give me the criteria, I will tell you how we measure success.
>
> I would measure the success of the program--which is not a single program, as you know, it is a multiplicity of programs--by the stated objectives of the successive pieces of legislation which now represent Federal policy.
>
> These build upon, as you are fully aware, the legislation of 1963, and even a superficial look at that structure of legislation would indicate that it has a variety of goals.
>
> If you look at the particular goals, you would have an opportunity to ask questions about whether they were realized or not.
>
> When it comes to the emphasis in the legislation upon preparation for entry into jobs, you have one set of measures that can be applied. If you raise a question about the distribution of Federal funds to produce leverage on new fronts in the vocational education enterprise so that the whole enterprise might be redirected, you have another set of measures. (p. 537)

Similarly, George Copa and Gary Forsberg of the Minnesota Research and Development Center for Vocational Education offered testimony that "All of the purposes of secondary vocational education must be considered when evaluating the effects of vocational education at this level; a process must be developed to specifically focus on the purpose of occupational exploration" (p. 401).

In sum, the testimony offered in the Congressional Hearings in September 1980 on current issues in vocational education seems overall to suggest three general conclusions. First, while the VEDS system seems generally to be considered an improvement over past similar evaluation efforts, it also was judged to be imposing a considerable data-gathering and financial burden on participating institutions (particularly postsecondary ones) and to be leading to the gathering of follow-up data on student employment success which were in the eyes of some of questionable validity and utility. Second, vocational education was generally considered to have a multiplicity of purposes including occupational exploration and preparation for advanced occupational preparation, as well as direct employment in vocations. Third, and in light of such a multiplicity of goals, there appears to have been considerable reluctance among witnesses to give clear priority to any one or two evaluation criteria for the purpose of evaluating the success of vocational education programs.

Abt Study of State and Local Compliance and Evaluation Practices (Beuke et al., 1980)

The December 1980 study by Abt Associates provides a third souce of information on the current status of evaluation practice with respect to vocational education. This investigation was based on "extensive on-site

interviews in 15 state departments of vocational education, 16 secondary education institutions, 12 postsecondary institutions, and 10 CETA prime sponsor organizations." These interviews were "supplemented by an analysis of extensive documents in each of the 15 states visited and a more limited set of documents from 41 additional states and territories" (p. 2).

The Abt inquiry looked into evaluation and data reporting requirements in light of the evaluation requirements of the 1976 Education Amendments, and in this regard concluded as follows:

> The most consistently implemented and according to the state agency staff the most important evaluation component in the 15 sample states is the program review process. . . . The major difficulty which states are encountering in meeting the requirements of evaluating planning and operational process (sic) is finding the resources needed to support the program review system. (p. 96)
>
> . . . . . .
>
> The next most consistently implemented evaluation category is student employment success, particular student follow-up--the category which was given greatest emphasis by Congress. . . . The problems which arise in meeting federal requirements for student follow-up data appear to stem from procedural difficulties of gathering accurate reliable data which are consistent across districts. The difficulty in gathering high quality data in turn results from poor response rates; problems in establishing standard definitions of the data elements together with the lack of conviction that standardized follow-up data are necessary; and from the lack of federal funding to ease the burden which the requirements impose on state and local resources.
>
> Formal employer follow-up data are collected much less consistently than student follow-up data . . . There is less agreement among vocational educators that employer follow-up, as opposed to student follow-up, is an important and necessary indication of vocational program effectiveness or that it can realistically be collected. (pp. 96-97)
>
> None of the states in our sample requires statewide reporting of student achievement data in all occupational areas, though a small number of states do have occupational proficiency standards in a limited number of occupations. (p. 97)

> Less attention has been given to evaluating the results of additional services to special needs populations. All sample states which have a formal program review process include some items on special needs populations, though the extensiveness of the review varies. In all cases, the review focuses on access to rather than results of vocational programs. (p. 98)

The Abt analysts also inquired into the implementation status of VEDS and reached the following conclusions:

> One of the major problems in implementing the reporting system has been the difficulty particularly at the post-secondary level, in arriving at common operational definitions of data elements . . . .
>
> A related problem is the accuracy of data entered into the system; people are counted in the wrong categories because of unclear definitions and local staff are forced to estimate figures under certain detailed breakdowns because local recordkeeping systems do not provide actual data for those categories.
>
> Another major problem with VEDS has been the substantial burden it has imposed on the state and local agencies responsible for collecting the required data. Meeting VEDS requirements requires an investment of substantial resources. The 1976 Amendments authorized funds to assist states in implementing VEDS, but the money was never appropriated and states have been forced to absorb the full costs of the system. (p. 101)

Wentling's Survey of State Evaluation Practices in Vocational Education, June 1981

A fourth and most recent source of information on the current status of vocational education programs is Wentling's (1981) survey. This survey, conducted in spring 1981, employed a survey questionnaire mailed to directors of vocational education in all fifty states and territories. Fifty of 57 questionnaires were returned (88% response rate). The four-page questionnaires sought answers to eight questions pertaining to vocational education evaluation, using checklist, rating and open-ended items. Wentling's data showed that state directors of vocational education rated the following items as ones they would rely on as indicators of program quality (shown in parenthesees are the number of state directors rating each item as an indicator):

Employer feedback (47)
Placement level (44)
Job satisfaction of grads (37)
Quality curriculum materials (37)
Instructor performance (36)
Student test performance (31)

Sentiments expressed regarding evaluation criteria appeared to be somewhat different in response to another question Wentling asked, namely:

> If federal requirements for evaluation were eliminated, what would you choose to evaluate?

Responses to the items listed were:

Placement (n=43)
Planning and operational processes (38)
Services to special populations (36)
Employer reaction (35)
Student performance (34)

Why services to special populations showed up relatively strongly as a preferred evaluation criterion in response to this question, but not with regard to the previously cited question regarding indicators of program quality is unclear.*

Despite these apparent preferences for evaluation criteria, it appears that actual reliance on them as evaluation criteria was not yet a fact of life in many states in the spring of 1981. For in another question Wentling asked respondents to indicate which evaluation activities were fully implemented in respondents' states. Responses were as follows:

- o Placement level of grads (35)
- o Employer assessment of grads (27)
- o Assessment of planning and operational procedures (21)

* One hypothesis to explain this anomaly is that apparently service to special populations was not included as a response option to the indicators question.

o Assessment of special needs services (18)

o Assessment of student performance (14)

Wentling's data also provide a clear indication of why preferences for evaluation were not fully implemented. For in response to the question, "What factors hampered the development of evaluation activities in your state?," the following items were each indicated by 10 or more respondents:

Inadequate resources (n=26)
Lack of federal technical assistance (20)
Lack of evaluation expertise (15)
Lack of guides and books (13)
Inadequate federal guidelines (12)
Negative attitudes toward evaluation by LEAs (10)

Moreover, when asked in an open-ended question what suggestions they might have for writing new legislation, the two most frequently offered suggestions were the following:

Provide more latitude for states to develop their own evaluation around their needs (n=9)

Provide funds for evaluation (n=7)

Also, in an apparent reflection of the strength of dissatisfaction in at least some quarters with regard to current evaluation requirements, one respondent suggested an action which would certainly be an unusual provision to include in new legislation, namely "Fire developers of guidelines."

In conclusion, Wentling offered five specific recommendations for the development of federal policy regarding vocational education evaluation:

1. Continue to emphasize the improvement function of evaluation and further develop ways of facilitating the use of results for improvement.

2. Analyze extant research findings and conduct new research on the validity of process measures for predicting products or outcomes of vocational education. These findings should provide the base for determining evaluative criteria.

3. The requirements for evaluation procedures should remain somewhat consistent to minimize state and local burden and to provide for the enhancement of currently used procedures.

4. Any new approaches should be tried out prior to the preparation and issuance of rules and regulations.

5. The use of evaluation results should receive added emphasis in terms of new procedures, staff development, and technical assistance. (p. 17)

And in a more general suggestion, Wentling advocated that "it is important to strive for a certain amount of stability of policy" with respect to evaluation of vocational education (p. 18).

* * * * * * *

These few paragraphs have briefly summarized recent evidence on the current status of evaluation of vocational education. The obvious next question is what should be done in the future with respect to federal requirements for evaluating vocational education? However, before addressing this question in the final section of this paper, let me make a brief detour to review other aspects of federal policy and experience in evaluating education programs.

IV. How Are Other Federal Education Programs Evaluated?

The reason for addressing this question is simple. Over the last fifteen years, a tremendous amount of effort has been invested in federally sponsored evaluations of education programs. While federal requirements for evaluating

other programs are not directly relevant to the evaluation of vocational education, a consideration of other programs and their evaluation requirements, in light of how they both resemble and differ from vocational education, will suggest some unique insights on vocational education evaluation policy.

There are, of course, many different federal education programs whose evaluation requirements might usefully be reviewed. Here, however, I only have time to consider two, namely, Title I of the Elementary and Secondary Education Act of 1965 as amended, and the Education for Handicapped Children program (PL 94-142). There are several reasons for choosing these programs as comparisons, but one obvious and direct reason is that these programs represent the major federal efforts to serve disadvantaged children and handicapped children, two of the special needs groups whom federal legislation explicitly mandates federally assisted vocational education programs to serve. Another very practical reason for considering these two programs is that they are the two federal education programs which probably reach directly into most of the LEAs also served by federal vocational education assistance, and therefore most local educative institutions which must meet federal requirements for evaluating vocational education programs also must meet the evaluation requirements for Title I and PL 94-142.

Title I

Title I constitutes the single largest federal source of revenue in support of public elementary and secondary schools. It provides money to counties and school districts to improve the educational achievement of educationally deprived children living in areas with high proportions of low income families. As mentioned already, Title I of 1965 was the first

major piece of federal legislation to include a provision for the regular evaluation of funded programs. McLaughlin's (1975) analysis of the legislative history of the evaluation provision in ESEA (PL 89-10) suggests two competing visions behind the evaluation mandate. To some, the evaluation provision was intended as a tool to provide parents with more information about program efforts funded under this title and hence to enhance the accountability of programs to the parents and communities they served. To others, particularly at the federal level, evaluation was viewed as a management and planning device for infusing rationality in subsequent program efforts. As a result there has been a continuing tension in Title I evaluation policy, similar to the tension identified by Hendrickson (1981) with respect to vocational education policy. Should Title I evaluation serve mainly the management and accountability needs of states, the federal administrative branch and the U.S. Congress, or should it be aimed more to serve the information needs of local officials, teachers and parents?

While the evaluation provision contained in PL 89-10 was a compromise, according to McLaughlin, intended to address both visions of the main ends that Title I evaluation was seen as serving, by 1967 the federal interest in evaluation as a management and planning tool clearly came to predominate in the educational amendments of that year (PL 90-247, section 404). With the 1967 amendments the Secretary of HEW was required to submit each year an evaluation report to all relevant Congressional committees, regarding the effectiveness of programs funded under Title I ESEA. The report was to include a detailed review and evaluation of programs for their entire past life, based to the maximum extent practicable on objective measurements. The clear intent was to use this information in directing future programs decisions at the federal level.

The federal need for evaluation information was further emphasized in the Educational Amendment of 1974 (PL 93-380) Section 151, which authorized the commissioner to provide for independent impact evaluation; to develop standards and evaluation models to be used by LEAs; to set objective criteria, and to develop a system that would produce data that are comparable statewide and nationwide, to provide technical assistance to LEAs in implementing the system; and to make an annual report to Congress on the evaluation results. While Section 151(a) of the 1974 amendments suggested the inclusion of parents' and project participants' opinions in any LEA evaluation, none of the rules and regulations that followed made any further mention of procedures for implementing this provision.

The Educational Amendments of 1978 (Section 124) saw a re-emergence of attention to the local perspective:

> (g) Evaluations.--A local educational agency may receive funds under this title only if--
>
> (1) effective procedures are adopted for evaluating at least every three years the effectiveness of the programs assisted under this title in meeting the special educational needs of educationally deprived children;
>
> (2) such evaluations will include, during each three-year period, the collection and analysis of data relating to the degree to which programs assisted under this title have achieved their goals, and will also include objective measurements of educational achievement in basic skills over at least a twelve-month period in order to determine whether regular school year programs have sustained effects over the summer; and
>
> (3) the evaluation will address the purposes of the program and that the results of the evaluations will be utilized in planning for and improving projects and activities carried out under this title in subsequent years (emphasis added).

By 1980, the mandate for the Commissioner of Education to develop standards and evaluation models for Title I evaluation had evolved into a complex system under which each LEA receiving Title I funds is required

to evaluate Title I programs at least once every three years, using reliable and valid instruments, procedures that minimize error, and yields a valid assessment of achievement gains. Toward the latter end the USOE sponsored the development of three alternative evaluation models

(the norm-referenced comparison group, and regression group models) for estimating the achievement gains of children served in Title I programs in the areas of reading, math and language. Each model is supposed to provide an estimate of the effect of children's receipt of Title I services in comparison to what participating childrem might have achieved in the absence of Title I services.* According to most recent evidence (Anderson et al., 1978) most LEAs are using the norm-referenced model in order to meet federal Title I evaluation requirements. Achievement gains estimates derived from LEA's applications of the three models (or at least in principle, any alternative model for estimating gains that is approved by the Commissioner) are to be reported to SEAs using a common metric, the normal curve equivalent or NCE. Each SEA can thus aggregate results of local Title I evaluations to report to federal officials who in turn can produce national level aggregations of the estimated impact of Title I nationwide. Because of the complexity of implementing the Title I evaluation requirements and models, the Education Department has instituted ten regional Technical Assistance Centers to help states and localities implement and improve Title I evaluations.

* It should be noted, however, that several evaluation experts and some empirical research has questioned the extent to which achievement gains estimates derived from these models are comparable.

Education for All Handicapped Children

In November 1975, the U.S. Congress passed the Education for All Handicapped Children Act (PL 94-142] mandating that by September, 1978, all school-aged handicapped children in the United States be assured "a free appropriate public education." Federal evaluation requirements implemented in connection with this act are markedly different than those pertaining to either Title I or vocational education, not only because the Handicapped Act is of more recent vintage than federal initiatives with respect to vocational education and education for disadvantaged children, but also because it embodies a quite different philosophy and approach to educational programming than either Title I or the various pieces of federal vocational education legislation.

The 1975 Education for All Handicapped Children Act specified a range of activities that schools must conduct with regard to handicapped children, activities that are not viewed as a unitary program, but instead are seen as procedural insurance that handicapped children will benefit from the rights they have been guaranteed under the Act and individually prescribed education.

> [The Act] requires that specialists be called upon to evaluate the children's special needs and determine the most appropriate educational environment for these children; that an individualized education program be developed fro each child identified as needing special education or related services; that the schools notify parents of findings concerning their children and include parents in the process of making decisions regarding how and in what circumstances their children will be educated; and that an opportunity for a hearing be provided to a parent who is dissatisfied with the schools' decision. Further, the Act asks that, to the extent that it is in the child's best interest, each handicapped child be educated with nonhandicapped children. (Bureau of Education for the Handicapped, 1979, p. 7]

According to the first annual report of the USOE's Bureau of Education for the Handicapped on implementation of the 1975 Act, federal contributions (under PL-94-142 Part B) amount nationwide to about 9% of total state and federal funding of education of handicapped children (Bureau of Education for the Handicapped, 1979, p. 214). Thus the federal share of funding for the handicapped appears to be roughly equal to the federal share of funding for vocational education -- according to the most recent data from NCES, federal expenditures provided about 8.5 percent of total vocational education expenditures in 1979 (U.S. Congress, 1980, p.89).*

So what then are the federal requirements for the evaluation of federally-aided education programs for handicapped children? To put the matter baldly, how do the federal evaluation strings attached to funds for education of the handicapped compare with the evaluation strings attached to federal funds for vocational education? The Boruch and Cordray study describes federal evaluation requirements for LEA's under PL-94-142 as follows:

> At the local level, the term evaluation refers primarily to diagnostic assessment of children. The regulations require that preplacement evaluation be conducted using multiple, appropriate assessment modes. If the child is found to have a handicapping condition, an Individualized Educational Plan (IEP) is devised. The content of the Indvidual Education Plan is required by the regulations to include: (1) an assessment of present levels of educational performance; (2) a statement of annual goals and short term instructional objectives; (3) a statement of specific special education and related services and an assessment of the extent to which the child is able to participate in regular education programs; (4) projected dates for inititiation and termination of services; (5) appropriate objective criteria, evaluation procedures and a schedule for reevaluation. (Boruch and Cordray, 1980, p. 3-10).

Obviously these evaluation requirements are quite different from

*In citing these data regarding federal contributions for the handicapped and vocational education, it should be noted that from the sources cited it is not altogether clear whether federal contributions as a share of total contributions have been calculated on exactly comparable bases. Thus the only conclusion that seems warranted is that federal contributions to education for the handicapped and to vocational education seem roughly comparable - on the order of 10% of total expenditures.

those mandated under either Title I or federal vocational education legislation. They pertain exclusively to procedural requirements and do not encompass any specific outcomes such as student achievement gains or employment success of handicapped children after leaving school. There appear to be two broad reasons for the difference: the first having to do with the assuumptions about the relationships between federal, state, and local levels of educational authority, and the second relating to assumptions about educational programming and children's individual educational needs.

In her account of the process of federal efforts to develop an evaluation plan for PL-94-142, Kennedy mentions this Act as unique because it "more clearly delineates the relationship between federal, state, and local agencies" than previous federal education legislation (Kennedy, 1978, p. 19). The essential idea was that states would have primary responsibility for implementing programs to insure children's rights under the law, and that the federal role was one of oversight responsibilities with respect to state activities. In the process of developing a plan for evaluating PL-94-142, a decision was reached that the primary audience for federal evaluation activities should be the "federal agency and the Congress" (Kennedy, 1978, p. 37).

Evaluation requirements in connection with PL-94-142 also were influenced by assumptions about the nature and appropriate level of decision-making regarding educational programming. One of the key ideas in PL-94-142 was the individualized education plan or IEP. Again Kennedy's account provides insight into some of the thinking behind this provision of PL-94-142:

> The Act implies a faith that those closest to children -- their teachers and parents will make the best decisions about children. The purpose of the Act is not to increase the academic or social growth of handicapped children but rather "to provide a free, appropriate public education" to all handicapped children.

> Most education programs have several goals, many of which may not even be related to the children -- for example increased efficienty in service delivery, fairness in assessment, equitable opportunities, improved parent satisfaction, or simply improved accountability. This is not to say that academic performance should be ignored but that evaluations of these aspects of educational programs may be more useful locally than nationally. The local area has more capability to modify and improve these aspects of programs which affect children's performance in educational settings (p. 25).

Apparently as a result of such thinking, federal evaluation requirements developed for PL-94-142 specify no particular outcome measures for the purpose of evaluating education programs for the handicapped. Instead federal requirements focus exclusively on procedural requirements; concerning diagnostic assessment, development of IEP's and guidelines on who should be involved in the development of IEP's.

* * * * * *

This brief review of federal evaluation requirements developed in connection with Title I and PL-94-142, and how they compare with requirements for the evaluation of vocational education, raises interesting questions. What is the appropriate balance between procedural requirements for the provision of educational opportunities, the attainment of certain outcomes in educational programs, and the consequences for individuals in life after schooling? What is the appropriate balance of activity and responsibility between the federal, state, and local levels of educational governance? There are I think no definitive answers to these questions, but this brief review of vocational education, education for disadvantaged children, and education for handicapped children should make it clear that the evaluation requirements connected with federal initiatives in these three areas are premised on some rather different answers to such questions. Why this is so

is a question worth scrutiny, I think, but for the present let me reiterate only that to the extent that local educational agencies meet the mandate for vocational education to serve disadvantaged and handicapped children they appear to be required to meet some quite different mandates for evaluating the services provided.

V. Federal Legislation: What Criteria Should Be Set Out for the Evaluation of Vocational Education?

As the discussion in the last section makes clear, answers to this question depend upon some rather fundamental assumptions about the appropriate balance of authority for vocational education among local, state, and federal levels of governance. While I have no answers beyond personal predilection regarding what such assumptions should be, I think it clear that when it comes to establishing evaluation criteria the federal role probably ought to be, given current circumstances, relatively minor. After all, the federal contribution to the funding of vocational education nationwide appears to be no more than 10% or so of the total effort. In this light, and because of the concerns of state vocational officials reviewed in section III, it seems reasonable that either the federal mandate for evaluation of vocational education should be lessened or federal funding should be provided to cover a substantial portion of the costs of meeting the federal evaluation mandate. At the same time, however, Wentling's advice, namely that whatever happens the federal government should strive for a certain degree of stability in terms of vocational education evaluation policy. After all the evaluation requirements set out in the Education Amendments of 1976 and elaborated in regulations in 1977 and in the 1979 policy memorandum are really only beginning to be implemented.

Purposes of Evaluation. One other general consideration concerning the purposes of evaluation is also worth setting out before discussing various types of evaluation criteria. As the review of the development of current evaluation requirements for vocational education (in section III), the brief discussion of evaluation requirements for Title I and PL-94-142 (in section IV) clearly suggest, one prime issue in the establishment and evolution of federal evaluation requirements has been the question of purpose: what ends and whose interests ought evaluation to be aimed at serving? In terms of federal mandates for evaluation there has been a continuing tension between what might be called national interests and local interests. On one hand evaluation has been intended to produce valid, accurate and comparable data, so that evaluation results can be aggregated from local to state levels, and from state to the national level to produce an overall view of particular programs. Of the three programs whose evaluation requirements have been reviewed Title I seems to have adhered most to this view of the purpose of evaluation.

Yet on the other hand evaluation requirements for all three programs (including as noted in section IV, increasingly in Title I) there has been the view that evaluation ought to provide locally relevant information which will be of assistance in improving programs at the local level regardless of whether or not the evaluation data are sufficiently standardized to allow aggregation beyond the local level. Among the three programs reviewed, this tendency seems to have been most preeminent with respect to PL-94-142.

What then should be the balance between these competing interests with respect to vocational education? Clearly it seems that the local program

improvement goal should have highest priority with respect to vocational education evaluation. First, as Hendrickson (1981) points out this seems to have been the predominant goal that Congress had in mind in writing evaluation requirements into the Educational Amendments of 1976 (even though, as Hendrickson also points out, this goal seems to have been somewhat slighted in the process of developing regulations for vocational education evaluation). Second, this view seems clearly to predominate among vocational education officials, as noted in the review of evidence on the current status of vocational education in section III. This suggests at a minimum, whatever criteria are established for the evaluation of vocational education in federal legislation, if their main purpose is to serve local program improvement efforts that they need not be implemented in a standard way so as to provide nationwide comparability and to allow aggregation acrosss the diverse types of institutions which provide vocational education programs at structurally different levels of the nation's educational system. In this regard, the Title I evaluation and reporting systems (TIERS) provides a sobering contrast with the Vocational Education Data System (VEDS). TIERS focuses essentially on three types of outcomes (namely reading, language, and math achievement) served by Title I programs in institutions which are relatively homogeneous (namely local education agencies). Yet even so, the effort to produce a national picture of Title I through TIERS has received some hard methodological criticism (see, for example, Jaeger, 1979; Linn, 1979; and Wiley, 1979). In contrast, consider how much more ambitious VEDS is than TIERS. Vocational education serves students in institutions of post-secondary education as well as ones governed by local education agencies. VEDS aims at gathering data not just on student achievement in basic academic skills, but also in occupational com-

petencies, and follow-up data on employment success defined in terms of placement and employer satisfaction. Moreover, despite the relative simplicity of TIERS in contrast to VEDS, the Department of Education has established a system of regional technical assistance centers (TACS) to help state and local education agencies implement the Title I evaluation and reporting system. No similar system has been established to provide technical assistance with respect to vocational education evaluation and reporting.*

In sum, then, it seems to me that there are several different grounds for concluding that the current federal mandate for evaluating vocational education is overly ambitious:

- In the testimony of state vocational education officials, the current requirements impose both reporting and financial burdens on agencies implementing vocational education programs with federal support.

- In comparison to federal evaluation and reporting requirements for Title I and PL-94-142, the requirements for vocational education are far more extensive, yet at the same time far less fully supported either in terms of funding or technical assistance.

- In the apparent intent of the U.S. Congress in writing evaluation requirements into the Education Amendments of 1976, and in the clear opinion of state vocational education officials, the major purpose of vocational education evaluation is program improvement at state, local, and institutional levels. Nonetheless, several of the evaluation criteria currently mandated for vocational education are of questionable utility in serving that end.

* It should be noted that the NCES does provide some limited support to state officials regarding the implementation of VEDS and vocational education evaluation. For example, NCES does provide financial support to cover travel costs for state officials to travel to other states to inspect their vocational education evaluation procedures.

Given this overall conclusion, what more can be said about specific types of evaluation requirements? Answers to this question are organized around the four types of evaluation criteria mentioned in the 1977 regulations and described in Table 1, namely, planning and operational processes; student achievement; employment success; and services to special populations. Since the employment success criteria were ones given special emphasis in the 1976 Amendments they are discussed first.

Employment Success

As mentioned in the introduction, the 1976 Education Amendments explicitly called for the evaluation of occupational vocational education in terms of the extent to which program completers and leavers find employment in occupations related to their training and are considered by their employers well-trained and prepared for employment. In the 1977 regulations the following measures were listed as appropriate indices of employment success:

(1) Rates of employment and unemployment;
(2) Wages rates;
(3) Duration of employment; and
(4) Employer satisfaction with performance of vocational education students as compared with performance of persons who have not had vocational education.

First note the differences between the legislative and regulatory language. Though the 1976 legislation mentions "employment ... related to ... training," the regulations speak more generally about rates of employment. I am not sure whether the shift in emphasis was intended or not. The literature on vocational education since 1976 still contains fairly frequent references to employment related to training. Nevertheless, de-emphasizing the "related to training" issue as part of a standard evaluation criterion seems eminently sound to me. Why? Because there are numerous different ways of determining job - training relatedness and different procedures can yield quite different results. While many follow-up studies on vocational education participants appear to rely on subjective judgements of job relatedness (usually as judged by teachers or former students) at least some studies rely on more systematic procedures for determining job relatedness to training. One study evaluated relatedness by systematically analyzing job titles in comparison to vocational education specialty area and found that there was only around 55% agreement between results of the systematic analysis and teachers judgements of job

relatedness.* This suggests clearly that basing evaluations of vocational education on subjective judgements of job relatedness places the endeavor on a very shaky foundation. While problems of subjectivity seemingly might be overcome by development and dissemination of standardized procedures for determining job-relatedness and for dealing with other practical problems as well, there are more fundamental problems, discussed below which suggest that this may not be a reasonable strategy.

A second type of indicator of employment success mentioned in both the 1976 legislation and in the 1977 regulations is employer satisfaction. Note, however, that there is a significant difference between the language of the two. The legislation referred to the extent to which former students are considered by their employers "to be well-trained and prepared," while the regulations mention "employer satisfaction with performance of vocational education students." I have no special insight as to how or why the legislative language was transformed into the regulatory language, though I can imagine several plausible reasons for the change -- for example, it might have been due to the supposition that while employers might have a hard time evaluating the previous training of their employees, they would be in a much better position to evaluate their employees' current job performance, performance which presumably reflects previous training. But whatever the cause for the switch in emphasis, it seems clear that it has caused some real problems. Apparently the focus in employer follow-up surveys on individual's work <u>performance</u> (rather than on their previous training) has caused concerns about implications for the privacy rights of individuals. According to the Abt report, some states have judged that privacy rights

*See Woods and Haney, 1981, section 4.5 for a fuller description of this study.

require that the permission of the former vocational education students and/or their parents be obtained before conducting follow-up surveys with the former students' current employers. This causes problems of both administrative burden and of inference. If some individuals do not give permission for their current employer to be surveyed nothing at all can be inferred about those individual's vocational education programs. According to the Abt report:

> One state in our sample feels that the need to obtain permission before contacting employers prohibits the administration of employer follow-up at the state level (Benke et al, 1980, p. 81).

Obviously such a conclusion, even in one state, must raise doubts about the appropriateness of evaluation criteria which cannot be used because of concerns for the privacy rights of former vocational education students.

Two other indicators of employment success were mentioned in the 1977 regulations, namely wage rates and duration of employment. It is easy to imagine why such criteria may have been suggested. Wage rates are, for example, one of the most widely recognized indicators of economic value. Indeed, classical economic theory would have us believe that wage rates in a freely competitive market are the only means available for comparing the value of different kinds of labor. And attention to duration of employment would seem to be a natural outgrowth of intention expressed in federal legislation as long ago as 1963 for vocational education to help individuals achieve employment stability (see p. 7). Yet using wage rates and employment duration as evaluation indices obviously raises numerous practical problems.*

Nevertheless, even putting aside the practical problems of relying

*With respect to wage rates, for example, the Department of Labor (1980) points out that "Earnings in an occupation ... vary by geographic region. ... In addition, workers in the same occupation may have different earnings depending on the industry in which they work" (p. 6). Such factors obviously make wage rates a highly imperfect indicator of the quality of occupational preparation.

on any of these indicators of employment success as criteria for evaluating the effectiveness of vocational education, it seems to me that there are two much larger problems with the entire logic behind reliance on any one or any one set of indicators of employement success. First, it is significant that substantial and apparently increasing proportions of secondary vocational education students do not cease their formal education upon completion of their high school careers. As noted already, according to follow-up data on the high school class of 1972, more than 50% of the students self-reported in their senior year of high school to be enrolled in a vocational program, undertake at least some form of postsecondary education within four years of high school graduation. Moreover, in the latest available report from NCES, focusing only on occupational vocational education at the secondary level, it was found that:

> Of students completing vocational program requirements, slightly more than half are available for immediate placement in the labor force. The proportion has dropped slightly since fiscal year 1972, from 58.9 to 55.2 in fiscal year 1978 (U.S. Congress, 1980, p.86).

If one is interested in indicators of employment success it is of course reasonable to focus attention only upon those who are in the labor force, or in the words of the NCES report are available for job placement. Indeed, such a delineation of attention with respect to measuring employment success is almost required in light of standard Department of Labor procedures for calculating indices such as labor force participation and unemployment rates.* Nevertheless, from the perspective of evaluating vocational education programs - that is activities designed to answer questions about what a program does in the interest of making judgements about the program - focusing

*In this regard, it should be noted that calculation of such commonly known economic indicators as unemployment rates is not as straightforward as might be assumed. In reanalyzing national longitudinal data sets, we have found (see Woods and Haney, 1981, chapter 4) that it is virtually impossible to reproduce DOL procedures precisely unless one has access to the same set of questions asked in DOL surveys.

heavily or exclusively on the employment success of former vocational education students tends to ignore the 40 to 50% of secondary occupational vocational education program completers who are not known to be immediately available for job placement. This point is especially crucial in light of the Congressional stipulation in the Education Amendments of 1976 that "in no case can pursuit of additional education or training by program completers or leavers be considered negatively in these evaluations" of programs purporting to impart entry level job skills (Sec. 112, b, 1).

This clearly suggests that follow-up studies on former vocational education students - including those in occupationally specific programs - ought to be construed more broadly than as efforts to determine the employment success of those in the labor force. Instead, in light of the apparent tendency for secondary vocational education students increasingly to pursue other activities (including further education, occupational training, and homemaking), it is clear that to afford a fuller picture of what vocational education programs do, follow-up studies should encompass a description of what program completeres and leavers are doing be they in or out of the labor force.

Despite this recommendation for broadening the scope of follow-up studies on the status of vocational education program completers and leavers, a note of caution should be interjected here regarding the likely utility of any kind of follow-up studies. As mentioned, the major purpose of vocational education evaluation - both in the apparent intent of Congress in writing the evaluation provisions of the 1976 Amendments and in the visions of vocational education officials surveyed in the studies summarized in section III - is <u>improvement</u> of vocational programs at the state, local, and institutional

levels. The problem in this regard is one of timing. Judging from previous evaluation reports, it seems not unreasonable to assume that follow-up studies on the post-program status of vocational education program completers and leavers take a full two years to complete - one year for individuals to enter the labor market or into whatever other activity they might undertake, and one year for the actual conduct and reporting of evaluation activities. Assuming that program managers do use evaluation findings to improve vocational program offerings, it seems reasonable to assume that another two years will elapse between evaluation reporting and the actual implementation of improvements based on those findings - one year for consideration and planning based on the findings and another year for developing and implementing improvements in terms of staff training, curriculum development, etc.*

Moreover, if the vocational education involved has a two year course of study, another two years will elapse before graduates from the revised improved program will graduate and be available for job placement, advanced training, or whatever. Thus it seems reasonable to assume that at least five years will elapse between gathering follow-up data and completion of the cycle of program improvement and graduation of students from the new improved program. Obviously then, there is a real limitation in this approach as a means of keeping vocational program offerings current with labor market trends, for labor markets can change substantially within a period of five years.

There is one final, and to my mind, most severe weakness in the idea

*This time projection is according to some recommendations, unrealistically short for program development and implementation. Weikart and Banet (1976), for example, argue that it can take as long as ten years to develop, test out, and implement new educational programs.

that follow-up surveys of the employment success of former vocational students will help make vocational program surveys responsive to labor market demands. Follow-up surveys can help to determine the employment success of completers and leavers of programs that are currently offered, but such surveys have little potential for identifying job openings and thereby potential program offerings that are new. In other words, follow-up surveys may provide data on the effectiveness of past program offerings, but have little potential for directly illuminating potential future offerings.

In arguing against the likely utility of follow-up employment success surveys as a means for keeping vocational education program offerings current with labor markets, I should note that I am not arguing against the general utility of follow-up surveys. Indeed, experience in educational research has clearly shown that follow-up surveys as long as 20 years after program completion can be of considerable value in influencing general thinking on education and general strategy of program offerings. Rather, I am simply noting the likely limited utility of follow-up surveys in meeting the shorter-term goal of keeping vocational program offerings current with labor market demands.

Student Achievement

As noted in the reviews of studies of vocational education program evaluation in section III, this appears to be the area in which provisions of the 1977 regulations are least fully implemented. Recall that in this area the regulations suggested reliance on indicators such as:

- standard occupational proficiency measures;
- criterion referenced tests; and
- other examinations of students' skills, knowledge, attitudes and readiness for entering employment successfully.

It is easy to understand why these provisions have not been widely implemented. Standard occupational proficiency measures simply do not exist for most occupations. Criterion referenced tests are a relatively new idea in the testing profession and the technology of criterion referenced testing is only beginning to be developed.* Moreover, there is considerable uncertainty about what factors are significant determinants (be they skills, knowledge, attitudes, connections, or whatever) of individuals' successful entry into employment. For example, one recent study of the labor market for male youth with less than a college degree argues that the youth labor market has structural features which are not widely understood, much less reflected in employment and training policy.#

Therefore as a gneral matter it would seem unwise for federal mandates to require any specific or standard approaches to measuring student achievement in occupational proficiencies. Nevertheless, there appears to be one area of student achievement that may be worthy of more attention, namely student achievement in the basic skills of reading, writing and mathematics. Imparting such skills is not traditionally seen as a priority goal of vocational education, but there are several reasons for concluding that more attention to basic skills achievement of vocational students is warranted. First, there has been widespread concern nationally over the last four years or so about

* See Berk, 1980, for an up-to-date assessment of the state of the art of criterion referenced measurement.

The study entitled Getting Started is by Paul Osterman (M.I.T. Press, 1980). While Osterman's work cannot be summarized here, let me note simply that the author argues, among many other points, that "much of the policy focus on improving the 'school to work transition' is misguided" (p. 154). He argues this because of evidence that male youth often undergo a period of employment instability, a period of settling down, and because "there is little relationship between the first job after school and later jobs" (p. 154).

the basic skills of all high school students, including college preparatory and general as well as vocational program students. This concern has been manifest most clearly in the minimum competency testing movement. Since 1975 more than half the states in the nation have initiated minimum competency testing programs for all public school students. Under such programs test results are used to determine award of high school diplomas, to control grade to grade promotion, and/or to identify students to receive remedial instruction. Most such programs schedule testing in the high school years, and test reading, math, and to only slightly lesser extent writing (Gorth, et al., 1980 for details). While many observers have questioned the likely efficacy of minimum competency testing as a strategy for improving learning, it seems undeniable that such testing reflects a widespread concern for the basic skills achievements of high school students. And as a result,all high school students including college prep, general, and vocation program enrollees (but in some places excluding handicapped students) will have to meet common standards of performance on basic skill tests.

Second, more attention to general basic skills seems warranted in light of the tendency documented above for substantial proportions of secondary vocational education program completers and leavers to pursue post-secondary education, additional training, or other activities prior to entering the labor force. In short, for the 50% or so of secondary vocational students who are not immediately available for job placement after leaving their vocational programs, basic skills may be a more general and relevant type of achievement to evaluate than specific occupational competencies.

Third, even for vocational students who go directly into the labor market, basic skills may be more important than occupationally specific com-

petencies. Why? One prominent economist, Lester Thurow, argues this point as follows. First he maintains that "most skills in the U.S. are not acquired in formal education or training; rather they are acquired through on the job training from one worker to another" (Thurow, 1979, pp. 324-325). He cites some "very clear" evidence to support the proposition. In light of this proposition, he argues that the prime purpose of vocational education ought to be creation of background characteristcs which tend to make individuals into low-cost trainees for potential employers. And from this, he argues third that "the three R's (reading, writing, and arithmetic) are an important ingredient in training costs" (p. 327). He also argues that good work habits are a second general type of characteristics which make individuals attractive to employers as potentially low-cost trainees. Also it should be noted that there are important qualifications to Thurow's general train of reasoning which for the sake of brevity I will not elaborate here. Nevertheless, it seems fair to summarize his general point as follows. Since employers tend not to select employees on the basis of their occupationally specific competencies, but instead on the basis of their trainability, basic skills attainment may be a more important goal for vocational education students planning to enter directly into the labor market than competencies keyed to specific jobs.

In arguing that basic skills attainment is worthy of more attention as an evaluation criterion for vocational education, I should make it clear what I am <u>not</u> arguing for. I am not arguing for any kind of national standard of basic skills attainment for vocational education students. While federal guidelines might, I think usefully direct more attention to evaluation criteria in this realm, I am not suggesting even that the federal government

make an effort to aggregate results of state and local assessment of basic skills into a national picture of the basic skills attainment of vocational students. The reasons for not suggesting a large federal role with respect to evaluation of basic skills attainment of vocational students are many, but here I mention only three. First, there seems to be widespread and widely accepted agreement to the proposition that federal, as opposed to national, standards of educational achievement would be an unwarranted breach of the long-standing division of authority over education in the United States. Second are the many technical problems that efforts to aggregate state and local evaluations of basic skills attainment would raise in trying to produce a national picture of vocational education.* Third, I see little likely utility from such an effort in terms of program improvement in light of the highly diversified and decentralized nature of the vocational education enterprise.

Services to Special Needs Populations

A third area for evaluation mentioned in the 1977 regulations is services to special needs groups, specifically:

> The results of additional services, as measured by the suggested criteria [pertaining to planning and operational procedures, student achievement, and employment success] ... that the state provides to these special populations:
>
> (1) Woman,
> (2) Members of minority groups,
> (3) Handicapped persons,
> (4) Disadvantaged persons, and
> (5) Persons of limited English speaking ability.

*See Jaeger, 1979; Linn, 1979; and Wiley, 1979 for a discussion of many of these issues in regard to efforts to aggregate local and state evaluations of basic skills effects estimates with respect to Title I programs.

As noted in the review of the Abt study and the Wentling survey in section III, evaluation activity regarding services to special populations appears to be relatively less than evaluation pertaining either to planning and operational procedures, or to employment success (though as noted, evidence in the Wentling study on the priority given to evaluation of services to special populations is somewhat ambiguous). As the Abt report noted:

> Less attention has been given to evaluating the results of additional services to special needs populations. All sample states which have a formal program review process include some items on special needs populations though the extensiveness of the review varies. In all cases, the review focuses on access to rather than results of vocational programs. (Beuke, 1980, p. 98).

In this regard, it seems to me that all that is reasonable to require in the way of regular evaluation of services to special populations are assessments of access to and participation in federally-aided vocational educational programs. To put the matter another way, the 1977 regulatory language requiring evaluation of the "results of additional services" provided to women, minority groups, handicapped and disadvantaged individuals, and persons of limited English speaking ability in terms of employment success and student achievement is unrealistic and unreasonable. I reach this conclusion for three broad sets of reasons.

First, are the general arguments already outlined regarding the limited support (either in terms of financial resources or technical assistance) with which the federal government has provided with respect to evaluation of employment success and student achievement effects of vocational education, and the limited utility of any such evalution in light of the apparent goal of federal evaluation requirements, namely program improvement at state, local, and institutional levels.

Second, is the language pertaining to results of additional services. If taken seriously, this would mean that special documentation would be required not only of general services provided to such special populations, but of additional services provided to each special population above and beyond the basic level of service provision which otherwise might be provided. This would mean that evaluation of the results of additional services would require addressing issues of supplemental services versus supplanted services - issues which have long bedeviled efforts to evaluate the Title I.*

Third, is the fact that most agencies providing federally-aided vocational education programs, namely local education agencies, already are under separate and quite different mandates to evaluate the results of service provision to precisely these same special populations. I have not had space in this brief paper to review the various federal mandates for evaluation of educational service provision to all the special populations mentioned in the 1977 vocational education regulations, but here let me only discuss the federal mandate for evaluation of educational programs for the handicapped, elaborated as a result of the passage of PL-94-142, which was briefly reviewed in section IV. As noted, federal evaluation requirements for LEA's providing services for the handicapped focus exclusively on procedural requirements (mostly in terms of individual educational plans) and do not specify attention to any particular outcome measures. Indeed enactment of federal legislation regarding evaluation for the handicapped (PL-94-142) seems to have

*Without trying to describe the history of the supplement versus supplant issue in Title I, and the various regulations that have grown up around this issue, it is worth noting that because of the ramifications of trying to identify supplemental services, the latest (1978) legislation provides for dropping of the supplemental requirement when Title I services are provided to students in districts serving a specified minimum proportion of disadvantaged students.

been premised on a notion that any a priori specification from the federal level of goals for handicapped students would be inappropriate, that the specific criteria for judging the effectiveness of individually prescribed educational plans for handicapped children should be specified at the local level. Thus, it seems to me that requiring the evaluation of vocational programs serving handicapped children in terms of any prespecified criteria of achievement and employment success would fly in the face of not only the procedural requirements of PL-92-142, but also the general philosophy which motivated the passage of this legislation by the U.S. Congress.

Planning and Operational Processes

The fourth area of evaluation mentioned in the regulations of 1977 was planning and operational procedures. Specifically mentioned under this rubric were the following sorts of characteristics of vocational eduation programs:

(1) Quality and availability of instructional offerings:
(2) Guidance counseling, and placement and follow-up services;
(3) Capacity and condition of facilities and equipment;
(4) Employer participation in cooperative programs of vocational education;
(5) Teacher/pupil ratios; and
(6) Teacher qualifications.

Though planning and operational procedures were mentioned in the 1977 regulations prior to the three other areas of evaluation mentioned (namely student achievement, employment success, and results of services to special populations), I have postponed discussion of this area of evaluation for the simple reason that I think there is greatest potential in this realm for specification of evaluation criteria which will most likely serve the intended goal of federal evaluation , namely program improvement. Why do I

give priority to evaluation of planning and operational procedures above student achievement, employment success, and results of additional services to special populations? The reason is fairly simple. Regardless of how good estimates of various program outcomes are, they provide absolutely no guidance for program improvement unless they are coupled with information on program processes. If one knows that program X helped to boost students' reading achievement by 10 points, this information is of little value if it is not accompanied by information on characteristics of the program (i.e. operational processes) which might help explain why it was effective. In other words, evaluation evidence on program outcomes such as achievement or employment success, however validly measured, provides virtually no basis for program improvement efforts if it is not accompanied by information on operational characteristics. Moreover, information on program characteristics, even if one does not have evidence on outcomes, still can provide a basis for improving programs; for example, if some independent standard exists for judging the characteristics which a program should have or hypotheses are available concerning likely relationships between operational program characteristics and program outcomes.

In this regard it seems to me that federal policy on vocational education evaluation might pay particular attention to documentation of the nature and extent of the vocational education enterprise. From the federal perspective, it seems likely to be far more useful to focus evaluation efforts on matters of vocational education program organization and administration, populations served and services provided, than on questions of effects, be they concerned with student achievement, employment success, or results for special populations.

There are two broad sets of reasons for advocating an evaluation focus on documentation of program nature and extent rather than on outcomes. This first has to do with what has been learned about systematic program evaluation over the last decade and a half and the second with the complexity of the vocational education enterprise and the federal role in it.

Since the mid-1960s systematic program evaluation has emerged as a major enterprise. Millions of dollars have been invested in it. Several new journals expressly intended to cover evaluation issues have been founded and evaluation appears to be emerging as a specialized profession, separate from research and administration. And as noted in section III, federal legislation on social programs now regularly contains provisions concerning evaluation of such programs. Nevertheless, not all has gone well for the evaluation enterprise. Systematic program evaluation has not usually been as useful, either in terms of policy-making or program management, as the early proponents of evaluation suggested (see Patton, 1978; Weiss, 1976, 1977). There are several different explanations for why systematic program evaluation has not proven to be as useful as might have been hoped, including, for example, the relatively primitive state of the art of evaluation, the failure to keep utilization in mind in designing and carrying out evaluations, and the myriad other influences besides evaluation findings on policymaking and program management (Lindblom & Cohen, 1979). Nevertheless, it seems clear that one reason why evaluations have not proven terribly useful is that there has been a tendency to focus on questions of effects or outcomes without paying sufficient attention to issues of specific purposes and characteristics of programs being evaluated. Such an approach has come to be known, usually derogatively, as the black-box model of evaluation. The main problems with

such an approach are two. First, it does little good to know that program X produced a gain of 10 points on an achievement test, or 10% greater chance of being employed unless such findings are accompanied by information on the nature of program X. Second, it is virtually impossible to produce effects estimates concerning any regular program which cannot be easily challenged by anyone with even a modicum of knowledge about technical evaluation issues. The problem is that without some kind of experimental manipulation of assignment of people to programs (for example, via random assignment), it is inevitably possible to hypothesize that estimated effects were due not to the program under study, but to the particular characteristics of individuals participating in the program. For this reason, Boruch and Cordray (1980) explicitly recommended "the use of randomized field experiments to plan and evaluate new programs, new variations of existing programs, and program components." There are, I think, strong grounds for questioning this recommendation, but it is nevertheless clear that for regular ongoing service programs such as vocational education, random assignment cannot (and in my view should not) be employed. This means that evaluations of the outcomes produced (the effects) of such programs can never be definitively assessed. This suggests to me that evaluation of vocational education programs should focus more on descriptive questions of what programs are and what people they serve rather than on the much more elusive question of what effects the programs have.* This is not meant to suggest that questions

* In this regard it is worth noting that as program evaluation has developed as a separate field of inquiry several groups have developed standards for program evaluation, and prominent among them are issues of program description. For example, the first standard set out by the Joint Committee on Standards for Educational Evaluation reads as follows:

> The object of the evaluation (program, product or project) should be described so that it is clear what form[s] of the object is [are] being evaluated. (Joint Committee, 1979)

Similarly, the number one guideline set out in the draft standards developed

of impact or effects are not of interest, but merely that given the last 15 years of experience with program evaluation, they cannot be very usefully addressed by mandating evaluation in terms of any one or two outcome criteria (nor for that matter any one or two sets of outcome criteria).

The second broad set of reasons for focusing evaluation on documentation of the nature and extent of programs rather than on outcomes has to do with the complexity of the vocational education enterprise and the limited federal contribution to that enterprise. As noted in section II, federally aided vocational education serves different kinds of goals (e.g., occupational exploration, preparation for advanced placement, and direct preparation for paid or unpaid employment), operates in many different institutions (28,000 at last count) and at different levels of the nation's educational system (secondary, postsecondary and adult), concerns dozens of different kinds of occupations, and is specifically mandated to serve seven different special populations. In contrast to the scope of the vocational education enterprise, however, federal aid to vocational education amounts to no more than 10% of total national expenditure on vocational education. In light of this contrast, and in light of the fact that the federal government has not provided significant support to states, local education agencies and

---

*(cont.)
by the evaluation research society is;

> The purposes and characteristics of the program or activity to be addressed in the evaluation effort should be specified as precisely as possible (cited in Boruch and Cordray, 1980, appendix).

institutions to meet the evaluation mandate in the 1976 legislation, much less the elaborated mandate in the 1977 regulation, it seems unreasonable for federal legislation to require evaluation in terms of any specific criteria or set of criteria.

What Else Can be Done to Make Evaluation Useful for Program Improvement?

Thus far I have argued mainly for curtailment of federal requirements for evaluation of vocational education outcomes. Nevertheless, the question of results is clearly of tremendous importance. The national interest obviously will not be well served by providing vocational education programs if those programs make no difference in the lives of individuals whom they aim at serving. Thus, in closing, it seems to be worthwhile to move beyond the fairly narrow question of evaluation criteria, and more broadly address the question of what else can be done to make evaluation useful for program improvements. While I do not have time nor space to elaborate much on this question, let me suggest five broad types of possibilities.

Evaluation Procedures. One strategy, apparent in the PL-94-142 legislation, would be to eschew the mandate of evaluation in terms of any specific criteria, and instead to place more emphasis on procedural requirements. Who should be involved in the conduct and reporting of evaluation, and who should have access to evaluation results? Instead of viewing former vocational education students (or in the jargon of current literature, program completers and leavers) only as objects of study, might evaluation policy seek to ascertain their views more directly on the strength and weaknesses of the programs in which they participated? Might not parents

of students, employers, and representatives of special populations, all of whose interests are intended to be served by federally aided vocational education, be systematically involved in evaluating this complex enterprise?

It should be noted that some such procedural requirements for evaluation already exist. As noted in the general description of federal requirements in appendix 1, for example, State Advisory Councils on Vocational Education already are mandated to participate in evaluations of vocational education at state levels, and the National Advisory Council on Vocational Education serves as the national counterpart to the state councils. Also, as noted in the Abt report (Beuke et al., 1980, p. 108), the Education Amendments of 1976 and 1978 mandate the establishment of three other sorts of groups at the state level (Committees for State Vocational Educational Planning, State Occupational Information Coordinating Committees, and State Employment and Training Councils) in order to coordinate CETA and vocational education programs. The establishment and operation of such coordinating groups has not been without problems, as the Abt report notes. But their existence (and according to the Abt report they "are having positive effects" Beuke et al., 1980, p. 123), clearly suggests that the locus of responsibility for vocational educational planning and policy-making resides more at the state than the federal level of government. If this is so, it would seem plausible to avoid direct federal mandate for evaluation in terms of any specific criteria, and instead to rely more on procedural guidance from the federal level.

Exemplary Program Identification. A second strategy would be to provide for incentives, beyond rules and regulations pertaining to specific evaluation criteria, to motivate useful evaluations of vocational education programs. One example of such a strategy is available in the Department of Education's Joint Dissemination Review Panel (JDRP). The JDRP regularly reviews evaluation evidence on the effectiveness of education programs in order to judge that evidence is persuasive. If a program is approved by the JDRP, it is listed in a Department of Education catalog, Education Programs That Work, and is eligible for federal support to promote dissemination of information about the program. The fifth edition of Education Programs That Work (1978) lists over 200 "exemplary educational programs," but among these only a single one is listed in the index under vocational education.* In contrast, exemplary programs listed under the rubrics of basic skills and early childhood education number more than two dozen each. There are several different hypotheses available to explain why vocational education programs show up relatively infrequently in this listing of exemplary educational programs,** but it

* It is the Vocational Reading Power project developed in Minneapolis, Minnesota, a state with one of the best developed systems for evaluating vocational education programs.

**Such hypotheses include: the apparent isolation in the federal bureaucracy of vocational education from other education programs, the relative prominence of state as opposed to federal authority in vocational education, the emphasis of the JDRP on"educational significance" as opposed to broader social and occupational significance, and greater federal investment in evaluating educational programs other than vocational programs.

seems to me very unlikely that it is because vocational education programs are generically less effective or exemplary than other educational programs. But whatever the cause, this fact suggests that vocational education has not exploited evaluation as a means of identifying good programs, publicizing information about them and thereby improving other programs, to as great an extent as have other federally aided education programs.

Technical Assistance. A third approach to evaluation which seems to me more likely beneficial in terms of helping to improve programs at state and local levels would be federal provision of technical assistance with respect to evaluation. To some extent this is of course already happening through activities of the National Center for Research in Vocational Education. This Center currently makes available a variety of materials concerning vocational education evaluation (for example, concerning the measurement of vocational education outcomes, and the characteristics of programs which tend to have high placement rates). Nevertheless, the federal activity in providing technical assistance with regard to Title I evaluations, through ten regional technical assistance centers, provides an interesting contrast. The model is of course not necessarily one to be directly emulated with respect to vocational education, but it is, I think, at least worth considering in an effort to identify ways in which federal provision of technical assistance with regard to evaluation can help states and local institutions use evaluation for the purpose of improving programs.

Evaluation Research into Special Issues. A fourth strategy for federal influence on evaluation of vocational education is targeted evaluation research on special issues. In this regard, the very complexity of the vocational education enterprise represents the various and sometimes competing interests involved in vocational education -- interests of federal, state, and local levels of government, interests of individual students and parents and those of diverse institutions, interests of business and organized labor, and various special interest groups as well. Current federal requirements for the evaluation of vocational education, it seems to me, clearly represent many of these different interests, and while I have argued against federally required evaluation criteria representing various special interests (e.g., the results of vocational education services provided to special populations), I am not arguing against the validity of the interests they represent. Rather I am simply arguing that it is needlessly burdensome (and I suspect highly inefficient as well in terms of deriving valid answers to specific questions) to require all institutions receiving federal vocational education funds to regularly evaluate their efforts in terms of such diverse criteria as currently mandated. Thus it seems to me that a more reasonable approach to deriving answers to specific questions (such as results for particular special populations) would be targeted evaluation research aimed at answering specific questions. Such a strategy could be easily coordinated with a federal approach to evaluation of vocational education which focuses on documentation and description of vocational education administration, services and populations. If such descriptive information

were available, it could be used to provide a framework in terms of which evaluation research targeted on specific questions could be organized, both in terms of elaborating specific questions and in terms of drawing systematic samples of vocational education programs and populations.

In suggesting the strategy of targeted evaluation research, I should again point out that the idea is not a new one. Indeed, the current vocational education study poses a prime example of such targeted research. Instead of relying exclusively upon the ongoing evaluation activities of the federal government, as in the Education Amendments of 1976, the U.S. Congress specifically mandated (and funded) this Study to address specific questions through "thorough evaluation and study" which the Congress wanted answered in preparation for reauthorization of federal legislation on vocational education. Also, the data now being acquired through VEDS provides an example of how descriptive data, acquired at the federal level, can help provide a framework for and thereby help refine evaluation questions addressed at the state and local levels. In this regard note that while NCES in reporting in 1980 on The Condition of Vocational Education drew a broad distinction between occupational and nonoccupational vocational education programs, officials from the state of Minnesota reported data showing that the amount of coursework students take in vocational education in Minnesota spans a continuum ranging from less than 100 hours to more than 800 hours (see U.S. Congress, p. 384). In other words, while the general distinction drawn at the federal level is a useful

one (far better than lumping together all secondary school students taking vocational coursework as simply "vocational education students"), a more detailed description developed at the state level revealed that such a dichotomization still conceals considerable variation in the extent of the vocational education training received by students.

Meeting Labor Market Demands. A last general issue pertains to the widespread interest in using evaluation to help improve the coordination of vocational education with labor market demands. As noted in section I, the widespread interest in making vocational education program offerings responsive to changing demands of the labor market is entirely understandable. It does little good to prepare individuals for occupations where there are no jobs available. Also, as noted, the concern for making vocational education more responsive to the labor market was clearly one of the prime motivations in the 1976 Congressional mandate to evaluate occupational vocational programs in terms of students' subsequent employment success (namely, placement in a related occupation, and employer satisfaction). Nevertheless, I also suggest that follow-up studies on employment success are not a terribly satisfactory means of serving the intended end -- both because of practical problems in conducting such follow-up studies while protecting the privacy rights of individuals, and more importantly, because such follow-up studies on former vocational education students inevitably will miss half the question. Specifically relying on follow-up studies will provide no direct information on occupational openings for which there are currently

no vocational program offerings. For example, in community Z, there may be a tremendous demand for computer programmers, but if there is no vocational program related to computer programming in that community, this fact would very likely not be revealed in follow-up studies.

If this is so, there remains the broad question of how mechanisms can be established to help make vocational education program offerings more responsive to labor market opportunities. The general answer, I think, is to bring other kinds of data to bear on vocational education planning and policy-making, namely surveys of current labor market opportunities and projections of future labor market demands. This idea is of course not a new one. It was reflected, for example, in the 1976 Congressional mandate to establish State Occupational Information Coordinating Committees, with membership to include state administrators of vocational education. Nevertheless, it should be acknowledged that the idea is not easy to implement for at least two broad sets of reasons. First, from a theoretical point of view, there is considerable uncertainty concerning market elasticities of skill substitution within or across occupations. As Thurow (1979) points out,

> The empirical magnitude of the elasticities of substitution depends on the level of aggregation of labor skills. At some levels of aggregation the elasticities are low; at others they are high. Generally, the more aggregate the vector of labor skills, the lower the elasticities of substitution between different skills. The elasticities of substitution between different types of electricians are presumably much higher than those between craftsmen and professional workers. (pp. 323-324)

In general Thurow seems to argue that elasticities of skill substitution are higher than typically recognized in vocational education and employment training circles. Nevertheless, at the same time he acknowledges that for some categories of skills, such as typing, elasticities of substitution may be low.

Yet leaving aside such theoretical problems, there is a more practical problem in attaining the goal of making vocational education more responsive to labor market demands. It is the fact that, whatever one wishes to assume about elasticities of skill substitution, there is a tremendous dearth of information systematically collected on labor market demands. Indeed, according to some recent accounts, the United States has the poorest system of collecting such information of any major industrialized country.

This suggests that meeting the goal of making vocational education more responsive to labor market demands would require the investment of considerable resources in developing a more comprehensive system for collecting, analyzing and reporting data on labor market demands. Such a task clearly seems beyond the ken of vocational education, or at least vocational education alone. Nevertheless, it seems to me that, however such a goal (namely the development of a better system for ascertaining present and future labor market demands) can be attained, at least one thing is clear. Labor market demand, as with other aspects of the economy, can fluctuate markedly across time and geography. According to the Department of Labor (1980):

> In every occupation and industry, the number of jobseekers and the number of job openings constantly changes. A rise or fall in the demand for a product or service affects the number of workers needed to produce it. New inventions and technological innovations create some jobs and eliminate others. Changes in the size or age distribution of the population, work attitudes, training opportunities, or retirement programs determine the number of workers available. As these forces interact in the labor market, some occupations experience a shortage of workers, some a surplus, some a balance between jobseekers and job openings. Methods used by economists to develop information on future occupational prospects differ, and judgments that go into any assessment of the future also differ.

This suggests that the task of coordinating vocational education programs offerings with current, much less future, labor market demand is not an easy endeavor. It also suggests to me that mandating individual education agencies and institutions to tailor their vocational education programs to labor market trends by way of their own follow-up surveys of former students is simply unreasonable. Instead I think a more reasonable strategy would be to invest additional resources in state or regional efforts to gather, analyze and report on labor market trends and to influence vocational education planning efforts by providing such information to local institutions.

REFERENCES

Berk, R. Criterion-referenced measurement: The state of the art. Baltimore, Md.: Johns Hopkins, 1980.

Beuke, V. et al. Implementation of the Education Amendments of 1976: A study of state and local compliance and evaluation practices. Cambridge, Mass.: Abt Associates, 1980.

Boruch, R., & Cordray, D. An appraisal of educational program evaluation: Federal, state, and local agencies. Evanston, Ill.: Northwestern University, 1980.

Bureau of Education for the Handicapped. Progress towards a Free Appropriate Public Education. (Report to Congress on the Implementation of Public Law 94-142: The Education for All Handicapped Children Act). Washington, D.C.: U.S. Office of Education, 1979.

Copa, G., & Forsberg, G. Measuring the employment and further education effects of secondary vocational education in Minnesota. In U.S. Congress, 1980, pp. 346-410.

Department of Labor. Occupational outlook handbook, 1980-81 edition. Washington, D.C.: U.S. Government Printing Office, 1980.

Education programs that work. A Resource of Exemplary Education Programs developed by Local School Districts and Approved by the Joint Dissemination Review Panel in the Education Division of the Department of Health, Education and Welfare (5th ed.). San Francisco: Far West Laboratory, for Educational Research and Development, 1978.

Gorth, W., et al. A study of minimum competency testing programs. Amherst, Mass.: National Evaluation Systems, 1980. (4 volumes)

Hendrickson, G. Evaluating vocational education: The federal stimulus. Washington, D.C.: National Institute of Education, 1981.

Jaeger, R. The effect of test selection on Title I project impact. Educational Evaluation and Policy Analysis, 1979, 1:2, 33-40.

Kennedy, M. Developing an evaluation plan for Public Law 94-142. In New Directions for Program Evaluation, Summer 1978, pp. 19-38.

Lindblom, C., & Cohen, D. Usable knowledge. New Haven, Cn.; Yale University Press, 1979.

Linn, R. Validity of inferences based on the proposed Title I evaluation models. Educational Evaluation and Policy Analysis, March-April 1979, 1:2, 23-32.

McLaughlin, M. Evaluation and Reform: The Elementary and Secondary Education Act of 1965, Title I. Cambridge, Mass.: Ballinger, 1975.

National Center for Educational Statistics. The condition of vocational education. In U.S. Congress, 1980, pp. 543-861.

National Institute of Education. The Vocational Education Study: The Interim Report. Washington, D.C.: Author, 1980. (Also reprinted in U.S. Congress, 1980, pp. 863-1168.)

Osterman, P. Getting Started. Cambridge, Mass.: MIT Press, 1980.

Oxford English Dictionary. New York: Oxford University Press, 1977.

Patton, M. Utilization-focused evaluation. Beverly Hills, Calif.: Sage, 1978.

Smith, E., & Holt, N. State of the art of vocational education evaluation: State evaluation procedures and practices. Belmont, Mass.: CRC Education and Human Development, 1979.

Thurow, L. Vocational education as a strategy for eliminating poverty. In The Planning Papers for the Vocational Education Study. Washington, D.C.: National Institute of Education, 1979, pp. 323-336.

U.S. Congress. Current issues in vocational education. Hearings before the House Subcommittee on Elementary, Secondary, and Vocational Education. Washington, D.C.: U.S. Government Printing Office 1980.

Weikart, D., & Banet, B. Model design problems in Follow Through." In A. Rivlin & M. Timpane (EDs.), Planned variation in Follow Through. Washington, D.C.: Brookings, 1976, pp. 61-77.

Weiss, C. (Ed.). The research utilization quandary. Policy Studies Journal, 1976, 4:3.

Weiss, C. Using social research in public policy making. Lexington, Mass.: Heath, 1977.

Wentling, T. A survey of state evaluation practices in vocational education. Mimeograph, University of Illinois, 1981.

Wiley, D. Evaluation by aggregation: Social and methodological biases. Educational Evaluation and Policy Analysis, March-April 1979, 1:2, 41-45.

Woods, E., & Haney, W. Proposed propositions and framework for study. Cambridge, Mass.; The Huron Institute, 1979.

Woods, E., & Haney, W. Does vocational education make a difference?: A review of previous research and reanalyses of national longitudinal data sets. Cambridge, Mass.: The Huron Institute, 1981.

Evaluation Requirements for Vocational Education

Funding for Federal Vocational Education programs is of two basic types: Formula grants to states and Discretionary grants. The evaluation process is different for each type. Here we only consider the evaluation requirements for the formula grants administered by the states.

State administered Vocational Education programs require evaluation at the state and federal levels. At the state level, formal evaluation is routinely conducted by two groups; the State Department of Vocational Education and the State Advisory Council on Vocational Education (SACVE). At the federal level, there is a parallel organizational scheme. The Bureau of Occupational and Adult Education (BOAE) within USOE and the National Advisory Council on Vocational Education (NACVE) serve as the federal level counterparts to the state agencies. The local administration of these programs is carried out by the district. The evaluation is typically informal, being composed of needs assessment and guidance regarding program operation provided by the Local Advisory Council on Vocational Education (LACVE).

(a) Evaluation Requirements at the State level. The law and regulations are explicit as to the content and procedures to be employed in the state evaluation. The evaluation is structured around a five-year program plan. The legislation explicitly states that the purpose of the evaluation is to revise and improve the programs conducted under this plan, this plan is jointly devised by representatives of the State Department of Education and the State Advisory Council (SACVE).

State Department of Education requirements. During the five-year period of the state plan, the State Department of Education is to evaluate the effectiveness of each program in terms of (a) planning and operational processes, (b) student achievement, (c) student employment success and (d) issues related to special populations. Further, the state is required to evaluate the extent to which individuals who complete or leave the program obtain employment in occupations related to their training and whether their employers consider them well-trained and prepared for employment. Sampling is permitted for this assessment. Finally, the State Department of Education is required to submit an annual accountability report which includes a description of how funds were used, a summary of the evaluations that were conducted and a description of how the evaluation information has been used to improve the state's program.

State Advisory Council requirements. Annually, the State Advisory Council is to prepare and submit to the Commissioner and National Advisory Council on Vocational Education, an evaluation report. Its contents are to include a synthesis of its evaluation of State Department administration and operation and the evaluations performed by the State Department of Education.

(b) Evaluation Requirements at the Federal level. An organizational structure, parallel to the state level, is established within the law for the federal level agencies. There are some notable differences in the explicitness of the evaluation requirements prescribed for the National Advisory Council, however.

Evaluation requirements for the Bureau of Occupational and Adult Education. At least ten states are to be reviewed during a given fiscal year. The purpose of the review is to analyze the strength and weaknesses of state programs. At the same time, DHEW is to conduct fiscal audits within those states. The Commissioner is to transmit to Congress a report on the National status of the Vocational Education programs. The report is to include information developed from the National Vocational Education Data System (VEDS), a summary of information obtained from federal reviews and audits and a synthesis of the evaluations performed by State Departments and State Advisory Councils.

Evaluation requirements for the National Advisory Council on Vocational Education. NACVE received a broadly stated evaluation function in the legislation. Its primary function is to provide policy-oriented annual reports and assessment of USOE/BOAE administration and operations.

Diversity in the Type of Evaluation Regulations

Examining the amount and type of information that is required across the four programs it is apparent that there are substantial differences. The direct grant type of program (e.g., Bilingual and the Discretionary grants for Vocational Education) have the least amount of oversight and reporting requirements. Title I and Vocational Education (Basic grants) are both state administered, formula allocation grants and have an additional level of evaluation imposed by the state agency. Vocational Education can be distinguished from Title I in that two agencies at the state and two agencies at the federal level are responsible for conducting routine evaluations. From this comparative assessment, we see that not only do the law and regulations indicate how evaluation is to be carried out, it can also influence how much is conducted and by whom.

APPENDIX 2

ESEA, Title I (Education of Disadvantaged Children, Basic Grants to LEAs)

(a) LEA Evaluation Requirements. The 1974 and 1978 Educational Amendments require the Commissioner to develop and make available to SEAs and LEAs (through the SEA) explicit standards and "models" for evaluation reporting at the local level. The October 12, 1979 Federal Register describes these standards and reporting regulations: every LEA receiving funding is required to submit an evaluation plan to the SEA that addresses how it will meet technical requirements of the regulation. At least once every three years, the LEA must evaluate its programs using "reliable and valid instruments," "procedures that minimize error" and a design that "yields a valid assessment of achivement gains." This latter requirement can be fulfilled by using one of three federally developed models or a suitable alternative approved by the SEA and Commissioner. Each model is supposed to provide an estimate of the effect of receiving Title I services on student performance compared to an estimate of what performance would have been in the absence of Title I services. Achievement scores are to be reported to the SEA using a common measure, a "normal curve equivalent" (NCE).

The new regulations also require longitudinal assessment to ascertain whether Title I gains are sustained after services are withdrawn. This assessment is for local use and reporting is not required unless requested by the SEA. Initial achievement status and gain, a description of the assessment process and project information are the only federally mandated evaluation requirements that are imposed on LEAs. The project information that is to be obtained includes: average duration of Title I service, pupil-per-teacher ratios, expenditures per child, and number of participants. According to the regulations, this project information is to be collected on a sample of grade levels.

(b) The SEA Evaluation Requirements. The SEA is charged with the responsibility for ensuring that the LEA educational plan is in compliance with the law and recently, this role has been expanded to include more extensive evaluation functions. SEAs are responsible for monitoring how the projects are carried out, providing technical assistance regarding LEA evaluation and aggregation of LEA data. The monitoring function is carried out through field visits by state Title I representative(s). The state receives one and one-half percent (set-aside) of its total allocation, or $150,000, whichever is greater, to perform these functions.

The SEA compiles the data that is submitted by the LEAs and submits (1) an annual performance report, containing: The number of participants served by types of service; number of participants by grade level for public and nonpublic recipients and "other information requested by the Commissioner" and (2) a biennial evaluation report, summarizing information for all or a representative sample of LEAs.

(c) Federal Evaluation Requirements. Section 183 of the 1978 Education Amendments clearly delineates the evaluation tasks and priorities to be addressed by the Commissioner. The law makes provision for two levels of evaluative evidence: independent evaluations designed to "describe and measure the impact of programs" and the Provision of Technical Assistance to States and local agencies on conducting evaluations. A maximum of one-half of 1 percent of the amount appropriated for these programs is provided for evaluation and priority is to be given to the federal assistance to state and local agencies.

APPENDIX 3

Public Law 94-142 (Education for the Handicapped) Evaluation Requirements

The Education for all Handicapped Children Act of 1975 and the pertinent regulations are explicit about responsibilities. The states are the primary target of federal oversight and they in turn are responsible for overseeing the local education agencies. The program is focused on the provision of a "free, appropriate public education for all handicapped children." The Bureau of Education for the Handicapped (BEH) in USOE was assigned the responsibility for administration and evaluation of P.L. 94-142.

(a) LEA "Evaluation" Requirements. At the local level, the term evaluation refers primarily to diagnostic assessment of children. The regulations require that preplacement evaluation be conducted using multiple, appropriate assessment modes. If the child is found to have a handicapping condition, an Individualized Educational Plan (IEP) is devised. The content of the Individual Education Plan is required by the regulations to include: (1) an assessment of present levels of educational performance; (2) a statement of annual goals and short term instructional objectives; (3) a statement of specific special education and related services and an assessment of the extent to which the child is able to participate in regular education programs; (4) projected dates for initiation and termination of services, (5) appropriate objective criteria, evaluation procedures and a schedule for reevaluation.

(b) SEA Evaluation requirements. The state has responsibility to ensure that the IEP has been prepared and that it meets the educational standards of the state. This is essentially a monitoring function and is carried out through on-site visits. Elaborate checklists have been developed by state agencies and BEH for assessing compliance with regulations. Additional monitoring requirements include fiscal audits and an assessment of the extent to which the Individual Educational Plan is actually carried out, in practice. This latter function is essentially a check to ensure that the program for individual children is actually implemented.

The law specifies that in any fiscal year, the state may use five percent of the total state allotment, under part B, or $200,000, whichever is greater for conducting required administrative activities. Evaluation in the sense of monitoring is included under this category of activities.

The State Education Agency is required to report (1) the number of handicapped children receiving services on October 1 and February 1 of the school year; (2) the number of handicapped children within each disability category; (3) the number of handicapped with each of three age groups. For all figures, unduplicated counts are required. This report is to be transmitted to the Commissioner.

(c) Evaluation requirements at the Federal Level. The Commissioner has responsibility for evaluation under Section 618 of the Act. Specifically, the legislation authorizes (1) annual studies; (2) assessment of the adequacy of information provided by state agencies; and (3) development of effective methods and procedures for evaluation.

This degree of explicitness is often not warranted for a variety of reasons. In some cases, the feasibility of conducting a specific type of evaluation is unknown and the details are deferred to the agency responsible for carrying out the evaluation. In other instances (e.g., ESEA, Title I) the law requires the development and implementation of models for estimating program effects but the SEA and LEA personnel are provided with some discretion as to which model they will follow.